

Appendix A1

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2006, 13:02:26 ; Search time 41 Seconds
(without alignments)
1072.465 Million cell updates/sec

Title: US-10-654-428-2
Perfect score: 2331
Sequence: 1 MARRTEPPDGGWGRVVLSA.....SLTAPGLLPLRLGHRRTVP 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	591	25.4	484	2	A55626
2	564	24.2	542	2	JC5507
3	555.5	23.8	494	2	A44458
4	546	23.4	494	2	JC4399
5	540	23.2	500	2	A55568
6	482.5	20.7	598	2	T32430
7	463.5	19.9	556	2	T16790
8	417.5	17.9	613	2	I39295
9	370.5	15.9	397	2	G69295
10	368.5	15.8	402	2	F95874
11	368.5	15.8	596	2	T23685
12	342	14.7	571	2	T33941
13	314.5	13.5	808	2	T16564
14	290.5	12.5	399	2	F83484
15	288	12.4	410	2	G82445
16	280	12.0	406	2	AG2704
17	280	12.0	414	2	G97486
18	250.5	10.7	407	2	AC1346
19	250.5	10.7	416	2	G69748
20	249.5	10.7	405	2	H81995
21	245.5	10.5	407	2	AF1716
22	245	10.5	418	2	F83986
23	241	10.3	426	2	TS0944
24	238.5	10.2	402	2	S47768
25	236	10.1	673	2	S59263
26	233	10.0	400	2	H91182
27	233	10.0	400	2	D86029
28	232.5	10.0	405	2	T44249
29	229.5	9.8	431	2	AB2949

30	229.5	9.8	433	2	H98333	hypothetical prote
31	228.5	9.8	398	2	AD0828	probable transmem
32	226.5	9.7	430	2	S74039	hypothetical prote
33	226	9.7	473	2	S38065	hypothetical prote
34	225	9.5	444	2	E83033	probable MFS trans
35	222	9.5	419	2	H83965	hypothetical prote
36	220	9.4	420	2	H69250	oxalate/formate an
37	218	9.4	388	2	AG3307	oxalate/formate an
38	208	8.9	373	2	S00656	hypothetical prote
39	206	8.8	408	2	H95184	hypothetical prote
40	205	8.8	408	2	H98051	oxalate/formate an
41	205	8.8	412	2	T37042	major facilitator,
42	201	8.6	501	2	S66816	probable nitrate/n
43	199.5	8.6	371	2	B90267	probable membrane
44	199.5	8.6	807	2	T20080	transport protein,
45	196.5	8.4	428	2	A75423	hypothetical prote

ALIGNMENTS

RESULT 1
A55626

monocarboxylate transporter MCT2 - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 18-Sep-1998

C:Accession: A55626

R:Garcia, C.K.; Brown, M.S.; Pathak, R.K.; Goldstein, J.L.
J. Biol. Chem. 270, 1843-1849, 1995

A:Title: cDNA cloning of MCT2, a second monocarboxylate transporter expressed in differer

A:Reference number: A55626; MUID:95130566; PMID:7829520

A:Accession: A55626

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-484 <GAR>

A:Cross-references: UNIPARC:UPI0000178E64; GB:L31957

C:Superfamily: monocarboxylate transporter MCT2

C:Keywords: transmembrane protein

QY	7	PPDGGWGRVVLSAFQSAIVFGVRSFGVFVEFAAFEEQAAARVSWIASIGIAVQFG	66
DB	14	PPDGGWGRVVLSAFQSAIVFGVRSFGVFVEFAAFEEQAAARVSWIASIGIAVQFG	73
QY	67	SPVGSALSTKFGPRPVMTGGLAALGMLASFATSLTHYLSIGLSSGSGWALTFAPTL	126
DB	74	GPISSVLVNNYGRPVVIGGLICTGMILASFSNSVELYLTIGFIGGLAFLNLPAL	133
QY	127	ACLSCEFSRRSLATGALTGVGLSFTFAPFQMLLSHYAMRGSLLVSAISLHVACG	186
DB	134	TIIGKYFYRRRPANGLAMAGSPVFLSSLAFFNQYLFNSYGMKGSFLILGIFLHSCVAG	193
QY	187	ALLRP-----PSLAEDPAVCGPRAQLT-----SLHNGPFLRYVAL	223
DB	194	CLMRPVQTSFPRKSKSKSVGRSDGSMKSKSVSTAERINRFLDFSLFRHGFLLYLSGN	253
QY	224	TLINTGYFIPYHLVAHLQDLWDPLPAFLISVAISDLVGRVSGWLGDA--VPGPVT	281
DB	254	VIMFLGFPAIFLAPYAKDKGVDEXNAALLSVMAFVDMFARPTGGLIANSKLIRPRIQ	313
QY	282	RLIMLWTLTGVSIALPVAQAPTAVALAVAYGFTSGALAPLAFSVLPGLIGTRIRYCG	341
DB	314	YFFSFALVGTGICHLCLPLADIVPALVVISIFFGYGFSGSVSLFETLMDLVGPARFSSA	373
QY	342	LGLLQMTESISIGLLGPLSGYLRDVSNGNYTASFVAVAGAFILSGSGLL	389
DB	374	VGLATIVECCPVLLGPPLAGKLVDKTKDYKMYIASGTLIVT-SGITL	420

RESULT 2

Hyphen 2 x A Z

JC5507

monocarboxylate transporter 3 - chicken

C/Species: Gallus gallus (chicken)

C/Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 09-Jul-2004

C/Accession: JC5507

R.Yoon, H.; Fanelli, A.; Grollman, E.F.; Philp, N.J.

Biochem. Biophys. Res. Commun. 234, 90-94, 1997

A/Title: Identification of a unique monocarboxylate transporter (MCT3) in retinal pigment

A/Reference number: JC5507; MUID:97312526; PMID:9168967

A/Accession: JC5507

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-542 <YOO>

A/Cross-references: UNIPROT:Q90632; UNIPARC:UPI0000171286; GB:U15685

A/Comment: This protein regulates lactate levels in the interphotoreceptor space.

C/Superfamily: monocarboxylate transporter MCT2

C/Keywords: phosphoprotein

F:141/Binding site: phosphate (Thr) (covalent) (by casein kinase II) #status predicted

F:210/Binding site: phosphate (Thr) (covalent) (by CAMP- and cGMP-dependent kinases) #st

F:234,467/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predict

F:262,526/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predict

Query Match

Best Local Similarity 24.2%; Score 564; DB 2; Length 542;

Matches 146; Conservative 82; Mismatches 185; Indels 66; Gaps 9;

QY 6 EPPDGGWGRVVTLSAFQSAALVGVLSFGVFVEFVAFAFEQARVSWIASIGIAVQOF 65

DB 16 KPPDGGWGWITVLFCHVITGFSYAFPKAVSVYFKELMKDFHVGYSSTAWISSIMLMLYG 75

QY 66 GSPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSGVALTFAPT 125

DB 76 TGPVCSIMVNOFCGRPVMLIGLLASSGMLASFNTNIIELYLTAQVLTGLGMLNFOPS 135

QY 126 LACISCFYSRRSLATGLATGVGLSSFTFAFPQWLLSHYAMRGSLLVLSLHLVAC 185

DB 136 LIMLGTVYDKRRPLANGLAAGSPVFLSLSPGLQVLEKFGWRGGLIMGLLNCCCTC 195

QY 186 GALLRP-----PSLAED-----PAVG-----GPRRA 205

DB 196 GAVVRPLDAGMKRTEKAQDKYEAKEMLMGCKSEEGISTTDGTTKTKAKKKPKKGL 255

QY 206 QLTSLHGHFRLRYVALTLINTGYFIPIHLVAHLQDLWDPLPAFLLSVAISDLVG 265

DB 256 LDFSIIFSNRGFIITISKFILVLGLFVPILLVNAYAKDTGVPDEAFAFLSIIGFIDIFA 315

QY 266 RVVSG-----WLGDAVPGPVTRLLMTTLTGVSALFPVAQAAPTALVALAVAGFTSG 319

DB 316 RPACGMVAGLKWVRPHV-AYLFSFAMLENGLTDICAR---ASNVTGLVIFCVFFGISYG 371

QY 320 ALAPLAFSVLPPELIGTRRYCGLLQMIIESIGLGPPLSGYLRDVSNGNYTASFVAGA 379

DB 372 MVGALQFEVLMATVGSQKFSASIGLVLLIEAFVILGPPSAGRLVDALKNYEVIFFYLAGS 431

QY 380 FLSSGSGILTLPHFFCFSTTSGPQDLVTEALPTKVPPLKEGLEGGINSTESGPESQS 438

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

DB 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

QY 432 EVVL-SALFLAMATYCCINRGKTP-----PPEKNPSAGGSDTE---EAS 474

A/Molecule type: nucleic acid

A/Residues: 1-494 <KIM>

A/Cross-references: UNIPARC:UPI000017072E; GB:M97382; NID:9191143; PIDN:AAB59630.1; PID:

A/Experimental source: CHO clone met-18b-2

A/Note: sequence extracted from NCBI backbone (NCBIP:118221)

R/Garcia, C.K.; Goldstein, J.L.; Pathak, R.K.; Anderson, R.G.W.; Brown, M.S.

Cell 76, 865-873, 1994

A/Title: Molecular characterization of a membrane transporter for lactate, pyruvate, and

A/Reference number: A49951; MUID:94170387; PMID:8124722

A/Accession: 148115

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-359, 'F', 361-494 <RES>

A/Cross-references: UNIPARC:UPI000012F3BF; GB:L25842; NID:9472332; PIDN:AAB59731.1; PID

C/Superfamily: monocarboxylate transporter MCT2

Query Match 23.8%; Score 555.5; DB 2; Length 494;

Best Local Similarity 31.8%; Pred. No. 2.5e-30;

Matches 135; Conservative 64; Mismatches 182; Indels 43; Gaps 5;

QY 7 PPDGGRVVTLSAFQSAALVGVLSFGVFVEFVAFAFEQARVSWIASIGIAVQOF 66

DB 13 PPDGGRVAVVGAFFISIGFSYAFPKSITVFFKEIGIFNATTSEVSWISSIMLAVMAG 72

QY 67 SPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSGVALTFAPT 126

DB 73 GPISSVLVNKKYGRPVMTAGCLSGCGLIASFNTVQELYLICIGVIGLGLAFNLPAL 132

QY 127 ACISCFYSRRSLATGLATGVGLSSFTFAFPQWLLSHYAMRGSLLVLSLHLVAC 186

DB 133 TWIGKYFYKRPPLANGLAAGSPVFLSTLAPLNQAFGIFGWRGSLILGILLNCCVAG 192

QY 187 ALLRP-----PSLA-----DPAVGPRQQLTSLH----- 212

DB 193 SLMRPIGPKPKIKELKSKESLQEAGKSEANTDLMGSPKGEKRSVLQITNKFLDLFLA 252

QY 213 HGFRLRYVALTLINTGYFIPIHLVAHLQDLWDPLPAFLLSVAISDLVGRVSGWL 272

DB 253 HRGFLLYLSGNVMEFFGLFTPLVFLSNYKSKQHSSEKSAFLLSILAFLVDMVARPSMGLA 312

QY 273 GDA--VPGPVTRLLMTTLTGVSALFPVAQAAPTALVALAVAGFTSGALAPLAFSVLP 330

DB 313 ANTKWIRPRIQYFPAASVANGVCHLAPLSTSYIGFCIYAGVGFACGWLSSVLFETLM 372

QY 331 ELIGTRRYCGLLQMIIESIGLGPPLSGYLRDVSNGNYTASF-----VAGAFILS 383

DB 373 DLVGPQRSSAVGLVTIVECCPVLLGPPLGLRLNDMYGDYKTYWACGVILLIAGIYLF 432

QY 384 GSGI 387

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

QY 433 GMGI 436

DB 433 GMGI 436

Appendix A3
us-10-654-428-2.rpr

F;60-81/Domain: transmembrane #status predicted <TM2>
F;87-105/Domain: transmembrane #status predicted <TM3>
F;112-136/Domain: transmembrane #status predicted <TM4>
F;145-172/Domain: transmembrane #status predicted <TM5>
F;176-195/Domain: transmembrane #status predicted <TM6>
F;196-255/Region: hydrophilic
F;256-280/Domain: transmembrane #status predicted <TM7>
F;292-314/Domain: transmembrane #status predicted <TM8>
F;322-342/Domain: transmembrane #status predicted <TM9>
F;346-368/Domain: transmembrane #status predicted <TM10>
F;380-402/Domain: transmembrane #status predicted <TM11>
F;416-438/Domain: transmembrane #status predicted <TM12>

Query Match 23.4%; Score 546; DB 2; Length 494;
Best Local Similarity 30.2%; Pred. No. 1.1e-29;
Matches 145; Conservative 66; Mismatches 215; Indels 54; Gaps 7;

QY 7 PPDGGRVVVLSAFQSAALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGIAVQOFG 66
Db 13 PPDGGRVAVVGAFAFISIGFSYAFPKSITVFKEIEIIFSAITSEVSWISSIMLAVMYAG 72
QY 67 SPVGSALSTKFGPRPVMTGILALGMLASFATSLTHLYLSIGLLSGSGWALTFAPTL 126
Db 73 GPISSTLVNKYGSRPVMIAGGCLSGCGLIAASFNTVOELVFCIGVIGGLAFNLNPAL 132
QY 127 ACLSCYFSRRSLATGLALTGVGLSSFTFAFFQWLLSHYAWRGSLLVASLSHLVACG 186
Db 133 TMIGKYFYKRRPLANGLAMAGSPVFLCTLAPLNOAFGIFGWRGSLILGILLNCVAG 192
QY 187 ALLRP-----PSLAE-----DPAVGSPRAQL-----TSLH 212
Db 193 SLMRPIGPQOGKVEKLKSKESLQKAGSDANTDLIGSPKGEKLSVFQTVNKFLLSLFT 252
QY 213 HGPFLRYTVALTLINTGYFIPYHLVAHLQDLWDPLPAFLLSVAISDLVGRVVSGL 272
Db 253 HRGFLLYSGNVVMFFGLFTPLVFLSNYKSKHFSSEKSAFLLSILAFVDMVARPSMGLA 312
QY 273 GDA--VPGPVTRLMLMTTLTGVSLLALFPVAQAPLVALVAAYGFTSGALAPLAFVLP 330
Db 313 ANTRWIRPRVOYFFAASVAVANGVCHLLAPLSTTYVGFCIYAGVFGFAFGWLLSVLEFETIM 372
QY 331 ELIGTRRIYCGLLQMLMESIGLLGPPLSGYLRDVSNGNYTASF-----VVAGAFLLS 383
Db 373 DLVGPQRFSSAVGLVTIVECCPVLLGPPLGLRLNDMYGDKYTYWACGVILLIAGLYLFI 432
QY 384 GSGI--LTLPHFFCFSTTTSGPQDLVTEALD-----TKVPLPKEGLEGLNSTES 432
Db 433 GMGINYRLVAKEQKAEEKRKDKEDDETSTDVDEKPKTKMETQSPAPLQNSSGDPAAEES 492

RESULT 5
A55568
monocarboxylate transporter 1 - human
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: A55568; T47157
R;Garcia, C.K.; Li, X.; Luna, J.; Francke, U.
A;Title: cDNA cloning of the human monocarboxylate transporter 1 and chromosomal localization
A;Reference number: A55568; MUID:95137602; PMID:7835905
A;Accession: A55568
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-500 <GAR>
A;Cross-references: UNIPROT:P53985; UNIPARC:UPI000016AD1B; GB:L31801; NID:G561721; PIDN:
A;Note: authors translated the codon GCA for residue 279 as Asp
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Well, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A;Reference number: 224379
A;Accession: T47157
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-479, 'A', 481-500 <AAA>

A;Cross-references: UNIPARC:UPI000012F3C0; EMBL:AL162079
A;Experimental source: adult melanoma (Memo cell line); clone DKFZp762B2310
C;Genetics:
A;Gene: GDB:SLC16A1; MCT
A;Cross-references: GDB:373211; OMIM:600682
A;Map position: 1p13.2-1p12
A;Note: DKFZp762B2310.1
C;Superfamily: monocarboxylate transporter MCT2

Query Match 23.2%; Score 540; DB 2; Length 500;
Best Local Similarity 30.6%; Pred. No. 2.8e-29;
Matches 133; Conservative 64; Mismatches 181; Indels 56; Gaps 6;

QY 7 PPDGGRVVVLSAFQSAALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGIAVQOFG 66
Db 13 PPDGGRVAVVGAFAFISIGFSYAFPKSITVFKEIEIIFSAITSEVSWISSIMLAVMYAG 72
QY 67 SPVGSALSTKFGPRPVMTGILALGMLASFATSLTHLYLSIGLLSGSGWALTFAPTL 126
Db 73 GPISSTLVNKYGSRIVMIVGGLSGCGLIAASFNTVOELVFCIGVIGGLAFNLNPAL 132
QY 127 ACLSCYFSRRSLATGLALTGVGLSSFTFAFFQWLLSHYAWRGSLLVASLSHLVACG 186
Db 133 TMIGKYFYKRRPLANGLAMAGSPVFLCTLAPLNOAFGIFGWRGSLILGILLNCVAG 192
QY 187 ALLRPSPLAEDPAVAG-----PRAQ----- 206
Db 193 ALMRP--IGPKPTAGKDKSKASLEKAGSGVKDLDHANTDLIGRHPKQEKRSVEQTIN 250
QY 207 ---LTSLHGPFLRYTVALTLINTGYFIPYHLVAHLQDLWDPLPAFLLSVAISD 262
Db 251 QFLDLTLFTHRG-FLLYLSGNVIMFFGLFAPLVLSSYKSKQHSSEKSAFLLSILAFVD 309
QY 263 LVGRVSGWLGDAVP--GPVTRLMLMTTLTGVSLLALFPVAQAPLVALVAAYGFTSGA 320
Db 310 MVARPSMGVANTKPIRPRIQYFFAASVAVANGVCHMLAPLSTTYVGFCVYAGFFGFAFGW 369
QY 321 LAPLAFSVLPBELIGTRRIYCGLLQMLMESIGLLGPPLSGYLRDVSNGNYTASF----- 374
Db 370 LSSVLFETLMDLVGPQRFSSAVGLVTIVECCPVLLGPPLGLRLNDMYGDKYTYWACGV 429
QY 375 -VVAGAFLLSGSGI 387
Db 430 LIISGYLFIQMG 443

RESULT 6
T32430
hypothetical protein C10E2.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32430
R;Wohlmann, P.; Sansone, J.
submitted to the EMBL Data Library, September 1997
A;Description: The sequence of C. elegans cosmid C10E2.
A;Reference number: Z21165
A;Accession: T32430
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-598 <WOH>
A;Cross-references: UNIPROT:O17328; UNIPARC:UPI000017B768; EMBL:AF026202; PIDN:AAB71245
A;Experimental source: strain Bristol N2; clone C10E2
C;Genetics:
A;Gene: CESP:C10E2.6
A;Map position: X
A;Introns: 53/3; 103/2; 145/2; 476/3; 569/3
Query Match 20.7%; Score 482.5; DB 2; Length 598;
Best Local Similarity 23.4%; Pred. No. 2.6e-25;
Matches 135; Conservative 75; Mismatches 181; Indels 185; Gaps 8;
QY 7 PPDGGRVVVLSAFQSAALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGIAVQOFG 66

Run on: September 7, 2006, 12:59:01 : Search time 201.3

(without alignments)
1404.426 Million cell updates/sec

Title:	US-10-654-428-2
Perfect score:	3337

Sequence: 1 MARRTPEPPDGGWGRVVTLSA

coring table: BLOSIM62

Gapor 10.0 , Gapext 0.5

searched: 2849598 seqs, 925015592 residues

total number of hits satisfying chosen parameters: 2849598

```

minimum DB seq length: 0
maximum DB seq length: 20000000000

```

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing.first 45 summaries

Database : UniProt 7.2:*

```
1: uniprot_sprot:*
2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Sult	No.	Score	Query Match	Length	DB	ID	Description
1	2151	92.3	426	2	Q7RTY0_HUMAN	Q7rtly0 homo sapien	
2	2142	91.9	424	2	Q2VP92_HUMAN	Q2vp92 homo sapien	
3	1899.5	81.5	428	2	Q66HE2_RAT	Q66he2 rattus norv	
4	1888.5	81.0	428	2	Q8CE94_MOUSE	Q8ce94 mus musculu	
5	1861.5	79.9	428	2	Q8CEH6_MOUSE	Q8ceh6 mus musculu	
6	1441.5	61.8	335	2	Q8CF46_MOUSE	Q8cf46 mus musculu	
7	1068.5	45.8	519	2	Q5M7D0_XENLA	Q5m7d0 xenopus lae	
8	938.5	40.3	447	2	Q5NC32_MOUSE	Q5nc32 mus musculu	
9	930.5	39.9	471	2	Q8NCK7_HUMAN	Q8nck7 homo sapien	
10	688	29.5	433	2	Q4KMI0_BRARE	Q4kmi0 brachydanio	
11	664.5	28.5	533	2	Q80IR0_XENLA	Q80ir0 xenopus lae	
12	645	27.7	492	1	MOT3_RAT	O70461 rattus norv	
13	643	27.6	290	2	Q9JTC0_MOUSE	O9jfc0 mus musculu	
14	640	27.5	486	2	Q8BGC3_MOUSE	Q8bgc3 m 14, 17 da	
15	637.5	27.1	463	2	Q5TFJ2_HUMAN	O5tfj2 homo sapien	
16	634.5	27.2	486	2	Q6ZSM3_HUMAN	O6zsm3 homo sapien	
17	632	27.1	460	2	Q6GMS9_XENLA	O6gms9 xenopus lae	
18	632	27.1	477	2	Q5O3M4_BRARE	O5o3m4 brachydanio	
19	631	27.1	492	1	MOT3_MOUSE	Q5b50 mus musculu	
20	631	27.1	492	2	Q5UB50_MOUSE	O5ub50 mus musculu	
21	629.5	27.0	504	1	MOT3_HUMAN	O95907 homo sapien	
22	624.5	26.8	473	2	Q6P2X9_XENTR	O6p2x9 xenopus tiro	
23	620	26.6	489	2	Q66HS9_RAT	O66hs9 rattus norv	
24	617	26.5	489	1	MOT2_RAT	O70451 mus musculu	
25	609	26.1	484	1	MOT2_MOUSE	O5rfh2 pongo pygma	
26	608.5	26.1	435	2	Q5RFH2_PONPY	O60669 homo sapien	
27	608	26.1	478	1	MOT2_HUMAN	Q3tcy8 mus musculu	
28	607.5	26.1	523	2	Q3TCY8_MOUSE	Q91w47 mus musculu	
29	607.5	26.1	523	2	Q91W47_MOUSE	Q8c086 mus musculu	
30	606.5	26.0	607	2	Q8C086_MOUSE	Q8nem3 homo sapien	
31	605.5	26.0	478	2	Q8NEM3_HUMAN		

ALIGNMENTS

Accession	Length	GC	Species	Gene
32	603.5	25.9	2	Q5RDA0_PONPY
33	601	25.8	2	Q4FK61_MOUSE
34	600.5	25.8	2	Q6PIX3_HUMAN
35	599.5	25.7	1	MOT7_HUMAN
36	599	25.7	2	Q7TMR7_RAT
37	595	25.5	2	MOT2_MESAU
38	594.5	25.5	1	Q4S087_TETTING
39	591.5	25.4	1	MOT4_HUMAN
40	591.5	25.4	2	Q5G591_HUMAN
41	591.5	25.4	2	Q2M1P8_HUMAN
42	583	25.0	2	Q6GM98_XENILA
43	581	24.9	1	MOT4_MOUSE
44	581	24.9	2	Q3UDF9_MOUSE
45	579.5	24.9	1	MOT4_RAT

RESULT 1

Q7RTY0_HUMAN
ID Q7RTY0_HUMAN PRELIMINARY; PRT; 426 AA.
AC Q7RTY0;
DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
DT 15-DEC-2003, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Monocarboxylate transporter 13.
GN Name=SLC16A13;
OS Homo sapiens (Human) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominida
OC Homo.
NCBI_TaxID=9606;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RP Halestrap A.P., Meredith D.;
RT "The SLC16 gene family - from monocarboxylate transporters (MCTs) to
RL aromatic amino acid transporters and beyond.";
PL Pflügers Arch. 0:0-0 (0).
[2]

1- MISCELLANEOUS: The sequence shown here is derived from an EMBL/GenBank/DBJ third party annotation (TPA) entry.

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>.
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; GN000145; CAD80156.1; -; mRNA.
ENSEMBL: ENSGN00000174227 ..

HGNC; HGNC:31037; SLC16A13.

GO; GO:0016021; C:integral to membrane; IEA

30; 30:0005315; E. + membrane; IEA.

GO; GO:0006810; P:transport; IEA

InterPro; IPR007114; MFS.

lileiPro; IPR011701; MFS_1.
Pfam: PF07690; MFS_1

PROSITE; PS50850; MFS: 1.

Membrane; Transmembrane; Transport.

SEQUENCE 426 AA; 44992 MW; 6A6849507C933624 CRC64

ry Match 92.38: Score 2151. 600

t Local Similarity	99.58%;	Pred. NO.	3.9e-143.
		DB 2;	

Conservative	1	Mismatches	1	Indels	0	Gaps	0
--------------	---	------------	---	--------	---	------	---

1 MARTEPPDGNRWWLSAFQSAlyGVIRSGVFVEFVAFEQARVSWIASIGI 60

1 MARTTEPDCGGMVWLSAFQSAIVFCYLRSGFVFVEFVAAFEECAARVSWITASTGT 60

61 AVOQFGSPVGSALSTKEGPRPVMTGCIALGMLASFATSI TH Y SI CI CCCCCC 10

61 AVQFGSPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSGSMAL 120
QY 121 TFAPTLACLSYFSRRSLATGLATGVGLSSFTFAPFQMLLSHYAMRGLLVASLST 180
DB 121 TFAPTLACLSYFSRRSLATGLATGVGLSSFTFAPFQMLLSHYAMRGLLVASLST 180
QY 181 HLVAAGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYLHLVAH 240
DB 181 HLVAAGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYLHLVAH 240
QY 241 LODLDWDPPLPAFLSVVAISDLVGRVSSGMLGDAVPGPVTRLMLMTTLTGVSALFPV 300
DB 241 LODLDWDPPLPAFLSVVAISDLVGRVSSGMLGDAVPGPVTRLMLMTTLTGVSALFPV 300
QY 301 AQAFTALVALAVAYGFTSGALAPLAFSVLPBLIGTRIRYCGLLQMIIESIGLGPPLS 360
DB 301 AQAFTALVALAVAYGFTSGALAPLAFSVLPBLIGTRIRYCGLLQMIIESIGLGPPLS 360
QY 361 GYLKDVSGNTASFTVAVAGAFLLSGSGLLTLPHFCESTTSGPQDLVTALDTKVPLPK 420
DB 361 GYLKDVSGNTASFTVAVAGAFLLSGSGLLTLPHFCESTTSGPQDLVTALDTKVPLPK 420
QY 421 EGLE 424
DB 421 EGLE 424
RESULT 2
QY 02VP92 HUMAN PRELIMINARY; PRT; 424 AA.
DB 02VP92 HUMAN PRELIMINARY; PRT; 424 AA.
QY 10-JAN-2006, integrated into UniProtKB/TrEMBL.
DB 10-JAN-2006, integrated into UniProtKB/TrEMBL.
QY 07-FEB-2006, entry version 3.
DB 07-FEB-2006, entry version 3.
QY SLC16A13 protein (Fragment).
DB SLC16A13 protein (Fragment).
QY Name=SLC16A13;
DB Name=SLC16A13;
QY Homo sapiens (Human).
DB Homo sapiens (Human).
QY Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DB Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
QY Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
DB Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
QY NCBI_TaxID=9606;
DB NCBI_TaxID=9606;
QY [1]
DB [1]
QY NUCLEOTIDE SEQUENCE.
DB NUCLEOTIDE SEQUENCE.
QY TISSUE=PCR rescued clones;
DB TISSUE=PCR rescued clones;
QY MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
DB MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
QY Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
DB Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
QY Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
DB Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
QY Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Hsieh F.,
DB Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Hsieh F.,
QY Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
DB Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
QY Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
DB Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
QY Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
DB Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
QY Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
DB Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
QY Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
DB Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
QY Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
DB Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
QY Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
DB Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
QY Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.R.A.,
DB Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.R.A.,
QY Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
DB Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
QY Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
DB Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
QY Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
DB Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
QY Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
DB Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
QY Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
DB Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
QY Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.,
DB Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.,
QY "Generation and initial analysis of more than 15,000 full-length human
DB "Generation and initial analysis of more than 15,000 full-length human
QY and mouse cDNA sequences."
DB and mouse cDNA sequences."
QY Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
DB Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
QY [2]
DB [2]
QY NUCLEOTIDE SEQUENCE.
DB NUCLEOTIDE SEQUENCE.
QY TISSUE=PCR rescued clones;
DB TISSUE=PCR rescued clones;
QY NIH MGC Project;
DB NIH MGC Project;
QY Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
DB Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
QY Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
DB Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
QY Distributed under the Creative Commons Attribution-NonDerivs license
DB Distributed under the Creative Commons Attribution-NonDerivs license

CC EMBL: BC109286; AA109287.1; -; mRNA.
DR NON TER 1
FT SEQUENCE 424 AA; 44789 MW; 9146D21BC6310E39 CRC64;
SQ
Query Match 91.9%; Score 2142; DB 2; Length 424;
Best Local Similarity 99.5%; Pred. No. 1.7e-142;
Matches 420; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 RTEPPDGMGRVVVLSAFQOSALVFGVLSFGVFVEFVAAFEEOA...VSWIASIGIAY 62
DB 1 RTEPPDGMGRVVVLSAFQOSALVFGVLSFGVFVEFVAAFEEOA...VSWIASIGIAY 60
QY 63 QQFGSPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSGSMALTF 122
DB 61 QQFGSPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSGSMALTF 120
QY 123 APTLACLSYFSRRSLATGLATGVGLSSFTFAPFQMLLSHYAMRGLLVASLSTLH 182
DB 121 APTLACLSYFSRRSLATGLATGVGLSSFTFAPFQMLLSHYAMRGLLVASLSTLH 180
QY 183 VACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYLHLVAHQ 242
DB 181 VACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYLHLVAHQ 240
QY 243 DLDWDPPLPAFLSVVAISDLVGRVSSGMLGDAVPGPVTRLMLMTTLTGVSALFPVAQ 302
DB 241 DLDWDPPLPAFLSVVAISDLVGRVSSGMLGDAVPGPVTRLMLMTTLTGVSALFPVAQ 300
QY 303 APTALVALAVAYGFTSGALAPLAFSVLPBLIGTRIRYCGLLQMIIESIGLGPPLSGY 362
DB 301 APTALVALAVAYGFTSGALAPLAFSVLPBLIGTRIRYCGLLQMIIESIGLGPPLSGY 360
QY 363 LRDVSGNTASFTVAVAGAFLLSGSGLLTLPHFCESTTSGPQDLVTALDTKVPLPK 422
DB 361 LRDVSGNTASFTVAVAGAFLLSGSGLLTLPHFCESTTSGPQDLVTALDTKVPLPK 420
QY 423 LE 424
DB 421 LE 422
RESULT 3
QY 06HE2 RAT PRELIMINARY; PRT; 428 AA.
DB 06HE2 RAT PRELIMINARY; PRT; 428 AA.
QY 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DB 11-OCT-2004, integrated into UniProtKB/TrEMBL.
QY 07-FEB-2006, entry version 18.
DB 07-FEB-2006, entry version 18.
QY Solute carrier family 16 (Monocarboxylic acid transporters), member
DB Solute carrier family 16 (Monocarboxylic acid transporters), member
QY 13.
DB 13.
QY Name=SLC16a13;
DB Name=SLC16a13;
QY Rattus norvegicus (Rat).
DB Rattus norvegicus (Rat).
QY Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DB Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
QY Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
DB Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
QY Muridae; Muridae; Murinae; Rattus.
DB Muridae; Muridae; Murinae; Rattus.
QY NCBI_TaxID=10116;
DB NCBI_TaxID=10116;
QY [1]
DB [1]
QY NUCLEOTIDE SEQUENCE.
DB NUCLEOTIDE SEQUENCE.
QY TISSUE=Kidney;
DB TISSUE=Kidney;
QY MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
DB MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
QY Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
DB Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
QY Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
DB Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
QY Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Hsieh F.,
DB Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Hsieh F.,
QY Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
DB Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
QY Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
DB Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
QY Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
DB Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
QY Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
DB Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
QY Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
DB Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
QY Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
DB Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
QY Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
DB Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
QY Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.R.A.,
DB Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.R.A.,
QY Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
DB Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2006, 12:59:01 ; Search time 301 Seconds
(without alignments)
1404.426 Million cell updates/sec

Title: US-10-654-428-2
Perfect score: 2331
Sequence: 1 MARRTEPPDGGWGRVVVLSA.....SLTAPGLLPLRLGHRRTVP 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapexc 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2151	92.3	426	Q7RTY0_HUMAN	Q7rtY0 homo sapien
2	2142	91.9	424	Q2VP92_HUMAN	Q2vp92 homo sapien
3	1899.5	81.5	428	Q66HE2_RAT	Q66he2 rattus norv
4	1888.5	81.0	428	Q8CE94_MOUSE	Q8ce94 mus musculu
5	1861.5	79.9	428	Q8CEH6_MOUSE	Q8ceH6 mus musculu
6	1441.5	61.8	335	Q8CF46_MOUSE	Q8cf46 mus musculu
7	1068.5	45.8	519	Q5M7D0_XENLA	Q5m7d0 xenopus lae
8	938.5	40.3	447	Q5NC32_MOUSE	Q5nc32 mus musculu
9	930.5	39.9	471	Q8NCK7_HUMAN	Q8nck7 homo sapien
10	688	29.5	433	Q4KM10_BRARE	Q4km10 brachydanio
11	664.5	28.5	533	Q801R0_XENLA	Q801r0 xenopus lae
12	645	27.7	492	MOT3_RAT	O70461 rattus norv
13	643	27.6	290	Q9JJC0_MOUSE	Q9jjc0 mus musculu
14	640	27.5	486	Q8BGC3_MOUSE	Q8bgc3 m 14, 17 da
15	637.5	27.3	463	Q5T7J2_HUMAN	Q5t7j2 homo sapien
16	634.5	27.2	486	Q6ZSM3_HUMAN	Q6zsm3 xenopus lae
17	632	27.1	460	Q6GM59_XENLA	Q6gm59 xenopus lae
18	632	27.1	477	Q503M4_BRARE	Q503m4 brachydanio
19	631	27.1	492	MOT3_MOUSE	Q503o8 mus musculu
20	631	27.1	492	Q5UB50_MOUSE	Q5ub50 mus musculu
21	629.5	27.0	504	MOT3_HUMAN	Q95907 homo sapien
22	624.5	26.8	473	Q6P2X9_XENTR	Q6p2x9 xenopus tro
23	620	26.6	489	Q66HS9_RAT	Q66hs9 rattus norv
24	617	26.5	489	MOT2_RAT	Q6j344 rattus norv
25	609	26.1	484	MOT2_MOUSE	O70451 mus musculu
26	608.5	26.1	435	Q5RFH2_PONPY	Q5rfh2 pongo pygma
27	608	26.1	478	MOT2_HUMAN	O60669 homo sapien
28	607.5	26.1	523	Q3TCY8_MOUSE	Q3tcy8 mus musculu
29	607.5	26.1	523	Q91W47_MOUSE	Q91w47 mus musculu
30	606.5	26.0	607	Q8C086_MOUSE	Q8c086 mus musculu
31	605.5	26.0	478	Q8NEM3_HUMAN	Q8nem3 homo sapien

32	603.5	25.9	465	2	Q5RDA0_PONPY	Q5rda0 pongo pygma
33	601	25.8	484	2	Q4FK61_MOUSE	Q4fk61 mus musculu
34	600.5	25.8	523	2	Q6P1X3_HUMAN	Q6p1x3 homo sapien
35	599.5	25.7	523	1	MOT7_HUMAN	O15403 homo sapien
36	599	25.7	522	1	Q7TMR7_RAT	Q7tmr7 rattus norv
37	595	25.5	484	1	MOT2_MESAU	P53988 mesocricetu
38	594.5	25.5	428	2	Q4S087_TERTNG	Q4s087 tetraodon n
39	591.5	25.4	465	1	MOT4_HUMAN	O15427 homo sapien
40	591.5	25.4	465	2	Q53G91_HUMAN	Q53g91 homo sapien
41	591.5	25.4	465	2	Q2M1P8_HUMAN	Q2m1p8 homo sapien
42	583	25.0	472	2	Q6GM98_XENLA	Q6gm98 xenopus lae
43	581	24.9	470	1	MOT4_MOUSE	P57787 mus musculu
44	581	24.9	470	2	Q3UDP9_MOUSE	Q3udp9 m bone mar
45	579.5	24.9	471	1	MOT4_RAT	O35910 rattus norv

ALIGNMENTS

RESULT 1									
ID	Q7RTY0_HUMAN	PRELIMINARY;	PRT;	426	AA.				
AC	Q7RTY0:								
DT	15-DEC-2003, integrated into UniProtKB/TrEMBL.								
DT	15-DEC-2003, sequence version 1.								
DT	07-FEB-2006, entry version 14.								
DE	Monocarboxylate transporter 13.								
GN	Name=SLC16A13;								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;								
OC	Homo.								
OX	NCBI_TaxID=9606;								
RN	[1]								
RP	NUCLEOTIDE SEQUENCE.								
RA	Halestrap A.P.;								
RT	"The SLC16 gene family - from monocarboxylate transporters (MCTs) to aromatic amino acid transporters and beyond.";								
RL	Pflugers Arch. 0:0-0(0).								
RN	[2]								
RP	NUCLEOTIDE SEQUENCE.								
RA	Halestrap A.P.;								
RL	Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.								
CC	-!- MISCELLANEOUS: The sequence shown here is derived from an								
CC	EMBL/GenBank/DBJ third party annotation (TPA) entry.								
CC	-----								
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms								
CC	Distributed under the Creative Commons Attribution-NoDerivs license								
CC	-----								
DR	EMBL; BN000145; CAD80156.1; -; mRNA.								
DR	Ensembl; ENSG00000174327; Homo sapiens.								
DR	HGNC; HGNC:31037; SLC16A13.								
DR	GO; GO:0016021; C:integral to membrane; IEA.								
DR	GO; GO:0016020; C:membrane; IEA.								
DR	GO; GO:0005215; F:transporter activity; IEA.								
DR	GO; GO:0006810; P:transport; IEA.								
DR	InterPro; IPR007114; MFS.								
DR	InterPro; IPR011701; MFS_1.								
DR	Pfam; PF07690; MFS_1.								
DR	PROSITE; PS50850; MFS; 1.								
KW	Membrane; Transmembrane; Transport.								
SQ	SEQUENCE 426 AA; 44992 MW; 6A6849507C933624 CRC64;								
Query Match									
		92.3%;	Score 2151;	DB 2;	Length 426;				
		Best Local Similarity	99.5%;	Pred. No. 3.9e-143;					
		Matches 422;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;			
QY	1 MARRTEPPDGGWGRVVVLSA	FFQSA	LVFGV	LSFGV	FFVEFVA	AEEQAA	RVSWIAS	IGI 60	
DB	1 MARRTEPPDGGWGRVVVLSA	FFQSA	LVFGV	LSFGV	FFVEFVA	AEEQAA	RVSWIAS	IGI 60	
QY	61 AVOQFGSPVGSALSTKFG	PRPVVMTG	GILALG	MLLAS	FATSLTH	LYLSIG	LLSGSGWAL	120	

Db 61 AVQFGSPVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYLSIGLLSGSGWAL 120
QY 121 TFAPTLACLSCYFSRRSLATGLALTGVGLSSFTFAPFQWMLSHYAWRGSLLVSALS 180
Db 121 TFAPTLACLSCYFSRRSLATGLALTGVGLSSFTFAPFQWMLSHYAWRGSLLVSALS 180
QY 181 HLVACGALLRPPSLAEDPAVGPRAOQLTSLHHGPFLRYTVALTLINTGYFIPYLHLVAH 240
Db 181 HLVACGALLRPPSLAEDPAVGPRAOQLTSLHHGPFLRYTVALTLINTGYFIPYLHLVAH 240
QY 241 LQDLWDPLPAFLSVVAISDLVGRVSWGMDAVPGPVTRLMLMTTLTGVSIALFPV 300
Db 241 LQDLWDPLPAFLSVVAISDLVGRVSWGMDAVPGPVTRLMLMTTLTGVSIALFPV 300
QY 301 AOPTALVALAVAYGFTSGALAPLAFSVLPBLIGTRRIYCGILQMIIESIGLLGPPLS 360
Db 301 AOPTALVALAVAYGFTSGALAPLAFSVLPBLIGTRRIYCGILQMIIESIGLLGPPLS 360
QY 361 GYLRDVSGNYTASFVVGAFLLSGGILLTLPHFCEFSTTSGPQDLVTEALDTKVPLPK 420
Db 361 GYLRDVSGNYTASFVVGAFLLSGGILLTLPHFCEFSTTSGPQDLVTEALDTKVPLPK 420
QY 421 EGLE 424
Db 421 EGLE 424

RESULT 2
Q2VP92 HUMAN
ID Q2VP92_HUMAN PRELIMINARY; PRT; 424 AA.
AC Q2VP92;
DT 10-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 10-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE SLC16A13 protein (Fragment).
GN Name=SLC16A13;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominda
OC Homo.
OC NCBI_TaxID=9606;
OX
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License

CC -----
DR EMBL; BC109286; AA109287.1; -, mRNA.
FT NON TER 1 1
SQ SEQUENCE 424 AA; 44789 MW; 9146D21BC6310E39 CRC64;
Query Match 91.9%; Score 2142; DB 2; Length 424;
Best Local Similarity 99.5%; Pred. No. 1.7e-142;
Matches 420; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 RTEPPDGMGRVVLSAFFQSALVGVLSFGVFEFVEFAAFEOAARVSWIASIGIAV 62
Db 1 RTEPPDGMVWVLLSAFFQSALVGVLSFGVFEFVEFAAFEOAARVSWIASIGIAV 60
QY 63 QQFGSPVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYLSIGLLSGSGWALT 122
Db 61 QQFGSPVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYLSIGLLSGSGWALT 120
QY 123 APTLACLSCYFSRRSLATGLALTGVGLSSFTFAPFQWMLSHYAWRGSLLVSALS 182
Db 121 APTLACLSCYFSRRSLATGLALTGVGLSSFTFAPFQWMLSHYAWRGSLLVSALS 180
QY 183 VACGALLRPPSLAEDPAVGPRAOQLTSLHHGPFLRYTVALTLINTGYFIPYLHLVAH 242
Db 181 VACGALLRPPSLAEDPAVGPRAOQLTSLHHGPFLRYTVALTLINTGYFIPYLHLVAH 240
QY 243 LLDWDPLPAFLSVVAISDLVGRVSWGMDAVPGPVTRLMLMTTLTGVSIALFPVA 302
Db 241 LLDWDPLPAFLSVVAISDLVGRVSWGMDAVPGPVTRLMLMTTLTGVSIALFPVA 300
QY 303 APTALVALAVAYGFTSGALAPLAFSVLPBLIGTRRIYCGILQMIIESIGLLGPPLSG 362
Db 301 APTALVALAVAYGFTSGALAPLAFSVLPBLIGTRRIYCGILQMIIESIGLLGPPLSG 360
QY 363 LRDVSGNYTASFVVGAFLLSGGILLTLPHFCEFSTTSGPQDLVTEALDTKVPLPK 422
Db 361 LRDVSGNYTASFVVGAFLLSGGILLTLPHFCEFSTTSGPQDLVTEALDTKVPLPK 420
QY 423 LE 424
Db 421 LE 422

RESULT 3
Q66HE2 RAT
ID Q66HE2_RAT PRELIMINARY; PRT; 428 AA.
AC Q66HE2;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Solute carrier family 16 (Monocarboxylic acid transporters), member
DE 13.
GN Name=Slc16a13;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG NIH MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; BC081902; AAH81902.1; -; mRNA.
DR Ensembl; ENSRNOC00000018785; Rattus norvegicus.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS50850; MFS; 1.
CW Membrane; Transmembrane; Transport.
IQ SEQUENCE 428 AA; 44993 MW; 7AC65745D23D3A79 CRC64;

```

Query Match      81.5%; Score 1899.5; DB 2; Length 428;
Best Local Similarity 88.2%; Pred. No. 2e-125;
Matches 372; Conservative 20; Mismatches 29; Indels 1; Gaps 1;

Y 1 MARRTEPDDGGRVVVLSAFQOSALVFGVLRSEGVFFVEFVAAFEEQARVSWIASIGI 60
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
b 1 MVHRTPEPDGGMVWLISAFQOSALVFGVLRSEGVFFVEFVAAFEEQARVSWIASIGI 60
Y 61 AVQOFGSPVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYSIGLSSGVAL 120
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
b 61 AVQOFGSPVGSALSTKLGPRPVMTGGILALGMLLASFATSLTHLYSIGLSSGVAL 120
Y 121 TEAPTTLACLSYFSRRSLATGLTGVGLSSFTPAEFFQWLLSHYAMRGSLIVSALS 180
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
b 121 TETPTLACLSRYFSQRRSLAMGLALTGVGISFAPAPLFQWLLNNYAMRGALLVSALS 180
Y 181 HLVACGALLRPPSLAEDPAVGCPRAQLTSLHGPFLRYTVALTLINTGYFIPYHLVAH 240
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
b 181 HLVACGALLRPLSTEDTAVGGPWAQITSLCHGPFLRYTVALTLINTGYFIPYHLVAH 240
Y 241 LODLDWDDLPAAFLLSVAISDLVGRVSWGMLGDVAPGPVTRLMLMTTLTGVSALFPV 300
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
b 241 LODLDWDDLPAAFLLSVAISDLVGRVSWGMLGDVAPGPVTRLMLMTTLTGVSALFPV 300
Y 301 AQAFTALVALAVAYGFTSGALAPLAFSVLPBELIGTRIRYCGGLQMIIESIGLLGPPLS 360
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
b 301 AQAFTTLVLAAYGFTSGALTVPAFSVIPELVGTGRIRYCGGLQMIIESVGLLGAPLS 360
Y 361 GYLDRVSGNYTASFVVAAGAFLLSGSGLLTLPHF-CRSTTSGPQDLVTEALDTKVPPLP 419
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
b 361 GYLDRVGTNYTASFVVAAGAFLLAGSGVLITLPHFSCISLSTSRPQDLVIEAPDTKIPPLP 420
Y 420 KE 421
  ||
b 421 KE 422

```

DT 07-FEB-2006, entry version 24.
DE 10 days neonate skin cDNA, RIKEN full-length enriched library,
DE clone:4732454C20 product:hypothetical protein, full insert sequence
DE (Solute carrier family 16 (Monocarboxylic acid transporters), member
DE 13) (Novel protein).
GN Name=Slc16a13; Synonyms=RP23-198E14.5; ORFNames=RP23-198E14.5-001;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aldinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Hummnick L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuura H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiuira K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Matick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fuke S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimomiya N.,
RA Nishio T., Okada M., Plessky C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RA RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;

Db 241 LODLGWDPLEPAFLLSVAASVDLVGRVASCWLGDAVPGPVARLIMLTTLTGVSLLALFPV 300
Qy 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPBELLIGTRRIYCGLLQMIIESIGLLGPPLS 360
Db 301 AQAASSTLVLAAYGFTSGALTTPVAFSVLPBELLVGTGRIRYCGMGLVQMIIESIGLLGAPLS 360
2y 361 GYLKRVSGNYTASFVAVAGAFLLSGGILLTLPHFF-CFSTTTGSPQDLVTEALDTKVPPLP 419
Db 361 GYLKRVYTGNYTASFVAVAGAFLLAGSGVLLTLPHFFSCISVSTSRPQDLVIEASDTKIPLP 420
2y 420 KE 421
Db 421 KE 422
RESULT 5
28CEH6_MOUSE
D Q8CEH6_MOUSE PRELIMINARY; PRT; 428 AA.
C Q8CEH6;
T 01-MAR-2003, integrated into UniProtKB/TrEMBL.
T 01-MAR-2003, sequence version 1.
T 07-FEB-2006, entry version 20.
E Adult male tongue cDNA, RIKEN full-length enriched library,
E clone:2310003J18 product:RIKEN CDNA 1700007D07.
N Name=Slc16a13;
S Mus musculus (Mouse).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
C Muridea; Muridae; Murinae; Mus.
X NCBI_TaxID=10090;
N [1]
P NUCLEOTIDE SEQUENCE.
C STRAIN=C57BL/6J; TISSUE=Tongue;
X MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
A Carninci P., Hayashizaki Y.;
T "High-efficiency full-length cDNA cloning.";
L Methods Enzymol. 303:19-44(1999).
N [2]
P NUCLEOTIDE SEQUENCE.
C STRAIN=C57BL/6J; TISSUE=Tongue;
X PubMed=16141072; DOI=10.1126/science.1112014;
A Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
A Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
Davis M.J., Wilmink L.G., Aidinis V., Allen J.E.,
Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
Kito H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
Liu S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
Schonbach C., Sekiguchi K., Sempile C.A., Seno S., Sessa L., Sheng Y.,
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
Sperling S., Stupka E., Sugiyama K., Sultana R., Takenaka Y., Taki K.,
Tammola K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yang J.,
Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
Wahlestedt C., Matic J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,

RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,
RT Hayashizaki Y.;
RL "The transcriptional landscape of the mammalian genome."; Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RT (Genome Network Core Team) and the FANTOM Consortium;
RL "Antisense Transcription in the Mammalian Transcriptome."; Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RT Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."; Nature 420:563-573(2002).
RL [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schonbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
RL [6]

NUCLEOTIDE SEQUENCE.
STRAIN=C57BL/6J; TISSUE=Tongue;
MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
"Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
Genome Res. 10:1617-1630(2000).
[7]
NUCLEOTIDE SEQUENCE.
STRAIN=C57BL/6J; TISSUE=Tongue;
MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
Sumi N., Ishi Y., Nakamura S., Hazama M., Nishine T., Harada A.,
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
"RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer."
Genome Res. 10:1757-1771(2000).
[8]
NUCLEOTIDE SEQUENCE.
STRAIN=C57BL/6J; TISSUE=Tongue;
Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
Katoch H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
Distributed under the Creative Commons Attribution-NonDerivs License

EMBL; AK028143; BAC25771.1; -; mRNA.
Ensembl; ENSMUSG0000044367; Mus musculus.
MGI; MGI:1916559; Slc16a13.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0016020; C:membrane; IEA.
GO; GO:0005215; F:transporter activity; IEA.
GO; GO:0006810; P:transport; IEA.
InterPro; IPR007114; MFS.
InterPro; IPR011701; MFS_1.
Pfam; PF07690; MFS_1; 1.
PROSITE; PS50850; MFS; 1.
Membrane; Transmembrane; Transport.
SEQUENCE 428 AA; 44878 MW; 83CFF037FA0407C5 CRC64;

Query Match 79.9%; Score 1861.5; DB 2; Length 428;
Best Local Similarity 86.0%; Pred. No. 9.5e-123;
Matches 363; Conservative 25; Mismatches 33; Indels 1; Gaps 1;

Db 181 HLMACGALLRPLSLTEDAAVGCGAQITSLLRHGPFLRYTAALTLINTGFIPYHLVAH 240
QY 241 LQDLMDPLPAFAFLSVVAISDLGVRVSWGIDAVGPEVTRLMLMTTLTGVSALFPV 300
Db 241 LQDLGMDPLPAFAFLSVAAVSVDLGRVASGWLGDVAPGVARLLMLMTTLTGVSALFPV 300
QY 301 AQAFTALVALAVAYGFTSGCALPLAFSVLPBELIGTRITXGGLQMIIESIGGLGPPLS 360
Db 301 AQAASSTLVLAVALAYGFTSGALTPLVAFSVLPBELVGTGRITXGGLVQMIIESIGGLGPPLS 360
QY 361 GYLDRVSGNYTASFVAVAGAFLLSGGILLTPHFF-CRSTTSGPODLVTEALDTKVPPL 419
Db 361 GYLRDVTGNYTASFVAVAGAFLLAGSGVLLTPHFFSCISVSTSRPQDLVIEASDTKIPPL 420
QY 420 KE 421
Db 421 KE 422

RESULT 6
Q8CF46_MOUSE PRELIMINARY; PRT; 335 AA.
ID Q8CF46_MOUSE
AC Q8CF46;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Adult male testis cDNA, RIKEN full-length enriched library,
clone:1700007D07 product:RIKEN cDNA 1700007D07.
GN Name=Slc16a13;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
Davis M.J., Wilming L.G., Aldinis V., Allen J.E.,
Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chaik A.M.,
Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
Jakt M., Kanapin L., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
Mottaqui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
Rost B., Ruan Y., Salzberg S.L., Sanderlin A., Schneider C.,
Schombach C., Sekiguchi K., Sempile C.A., Seno S., Sessa L., Sheng Y.,
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
Sperling S., Stupka E., Sugitara K., Sultana R., Takenaka Y., Taki K.,
Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,

RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmer S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimiya N.,
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Ohno Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaio I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gajobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Brad D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan A., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gajobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaio I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Kashiwagi K.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Koyama S., Kurihara C.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AK005699; BAC25120.1; -; mRNA.
DR Ensembl; ENSMUSG0000044367; Mus musculus.
DR MGI; MGI:1916559; Slc16a13.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1.
DR PROSITE; PS50850; MFS; 1.
DR Membrane; Transmembrane; Transport.
DR KW SEQUENCE 335 AA; 35061 MW; 087007B8A17B7DCF CRC64;
SQ
Query Match 61.8%; Score 1441.5; DB 2; Length 335;
Best Local Similarity 84.5%; Pred. No. 2.7e-93;
Matches 278; Conservative 24; Mismatches 26; Indels 1; Gaps 1;
QY 94 MLASPAATSLTHLYSIGLSSGVALTFAPTLACISCFRRRSIATGLTGVGLSSF 153
DB 1 MLASPAATSLTHLYNIGLSSGVALTFTPTMACLSRYSQROSLAMGLALTGVGISSF 60
QY 154 TFAPEQWLISHYAWRGSLIVSALSLHIVACGALRPPSLAEDPAVGGRPRAQLTSLHH 213
DB 61 AFAPLFQWLINNYAWRGALIVSALSLHMACGALRPLSLTEDAIVGPGAQITSLRH 120
QY 214 GPFRLYVALLTINTGYFIPYLHVAHLQDLWDPLPAFALLSVVAISDLVGRVSGWLG 273
||||| ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||

b 121 GPFLLRYTAAALTLINTGFEFIPYVHLVAHLQDLGWDPPLPAFLLSVAAVSDLVGRVASGWLG 180
Y 274 DAVPGPVTRLMLMTTLTGVSIALFPVAQAPALVALAVAYGFTSGALAPLAFSVLPPELI 333
b 181 DAVPGPVARLMLMTTLTGVSIALFPVAQASSTLVVALAVAYGFTSGALTPVAFSVLPPELV 240
Y 334 CTRRIYCGLGLQMIIESIGLLGPPLSGYLRDVSGNYTASFVAVAGAFLLSGSGLLTLP 393
b 241 GTGRITCCGMGLVQMIIESIGLLGAPLSGYLRDVTGNNTASFVAVAGAFLLAGSVLLTLP 300
Y 394 FF-CFSTTSGPDQDLVTEALDTKVPLPKE 421
b 301 FFSCISVSTSRPDQDLVIEASDTKIPLPKE 329

ESULT 7
5M7D0 XENLA
D QSM7D0 XENLA PRELIMINARY; PRT; 519 AA.
C QSM7D0;
T 01-FEB-2005, integrated into UniProtKB/TrEMBL.
T 01-FEB-2005, sequence version 1.
T 21-FEB-2006, entry version 14.
E LOC496238 protein.
N Name=LOC496238;
S Xenopus laevis (African clawed frog).
K Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
K Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
K Xenopodinae; Xenopus; Xenopus.
K NCBI_Taxid=8355;
N [1]
P NUCLEOTIDE SEQUENCE.
C TISSUE=Testis;
X MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
A Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
A Richardson P.;
T "Genetic and genomic tools for Xenopus research: The NIH Xenopus
T initiative.";
T Dev. Dyn. 225:384-391(2002).
N [2]
P NUCLEOTIDE SEQUENCE.
C TISSUE=Testis;
X MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
A Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
A Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
A Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
A Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
A Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
A Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
A Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
A Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
A Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
A Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
A Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
A Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
A Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
A Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
A Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
A Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
T "Generation and initial analysis of more than 15,000 full-length human
T and mouse cDNA sequences.";
T Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
N [3]
P NUCLEOTIDE SEQUENCE.
C TISSUE=Testis;
A Klein S., Gerhard D.S.;
N Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
C -1- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein (By
C similarity).

DR EMBL; BC088702; AAH8702.1; -, mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS50850; MFS; 1.
KW Membrane; Transmembrane; Transport.
SQ SEQUENCE 519 AA; 55329 MW; 0B438988ECC681F8 CRC64;

Query Match 45.8%; Score 1068.5; DB 2; Length 519;
Best Local Similarity 53.0%; Pred. No. 7.9e-67;
Matches 227; Conservative 56; Mismatches 120; Indels 25; Gaps 5;

QY 12 MGRVVLSAFQSAIVFGVLRSGVFVEFVAAFEQARVSMIASIGIAVQQFGSPVGS 71
Db 12 MGVVLLLSCFLLQAGLVFGTVRSFGVFFPEFVEIFESAGSVSWTSCGVAVQQLMSPLGC 71

QY 72 ALSTKFGPRPVMTGCIILALGMLLASFATSLTHYISIGLLSGSGWALTFAPTLACLSC 131
Db 72 SLAQRFGSRPVVILGLMSALGMFFASFATELYQLYLCIGLSGLGMALIFSPITMAVTR 131

QY 132 YFSRRSLATGLALTGVLSFTFAPFQWLLSHYARSGSLLVLSALSLHVACGALIRP 191
Db 132 SFVRRSLATGFLMVGVSFALSPLLQYLLLETYSRGALLLSGLALHSVPCGALLR- 190

QY 192 PSLAEDPAVGPRPRLTS-----LHNGPFLRYTVALTLINTGYFIPYHLVAHLQDL 245
Db 191 -----PLPGDPRESSYSFEMGWKLLGEGVFLRYCLAITSINTGYFVPYHLVAHMQARG 244

QY 246 WDPLPAFLLSVVAISDLVGRVSGWLGDAVPGPVTRLMLMTTLTGVSIALFPVAQAPT 305
Db 245 AGEWHAALIMSVIGADVGRLLFAGWLSDLAPQRTLLFLSLWTALTGVILGLPLAKEV 304

QY 306 ALVALAVAYGFTSGALAPLAFSVLPPELIGTRRIYCGLGLQMIIESIGLLGPPLSGYLRD 365
Db 305 AMGAALAVAFGCAGALTPGVFSALPGVVDQRLPALGLQMIIESGGGLGAPLSGWLCD 364

QY 366 VSGNYTASFVAVAGFLLSGSGLLTLPHPFCFSTTSGPDQDLVTEALDTKVPLPKELEG 425
Db 365 LTGDFTLSPAVAGSFLTLGALLTLTPRVFCNKQT-----DL---SLKDKAEGPPQGSWD 416

QY 426 GLNSTESG 433
Db 417 G-----ESG 420

RESULT 8
Q5NC32 MOUSE
ID Q5NC32_MOUSE PRELIMINARY; PRT; 447 AA.
AC Q5NC32;
DT 01-FEB-2005, integrated into UniProtKB/TrEMBL.
DT 01-FEB-2005, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Novel RNA polymerase Rpb1 C-terminal repeat domain containing protein.
GN Name=RP23-198E14.4; ORFNames=RP23-198E14.4-001;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sycamore N.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC EMBL; AL669869; CAI35220.1; -, Genomic DNA.
DR Ensembl; ENSMUSG00000040938; Mus musculus.

DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1.
DR PROSITE; PS50850; MFS; 1.
KW Membrane; Transmembrane; Transport.
SQ SEQUENCE 447 AA; 45235 MW; DEDD0C70314FAFD8 CRC64;

Query Match 40.3%; Score 938.5; DB 2; Length 447;
Best Local Similarity 46.1%; Pred. No. 9.7e-58;
Matches 202; Conservative 70; Mismatches 147; Indels 19; Gaps 6;
OY 7 PPDGGRVVVLSAFQSAIVFGLRSFGVFVEFVAAFEEOAARVSWIASIGIAVQOFG 66
DB 8 PPDGGRVVVLSAFVAVNGLSYGLRLSLGLALPDLAHFERSAQDTAWVSALALAVQOAA 67
OY 67 SPVGSALSTKFGPRPVMTGILALGMLLASFATSLTHLYLSIGLSSGWSALTFAPTL 126
DB 68 SPVGSALSTRWGARPVMVGVLTSLGLVSAFARSLHLHYLGLGLAGSGWALVFAPAL 127
OY 127 ACLSCYFSRRSLATGLALTGVGLSFTFAPDFQWLSHYAWRGSLLVASLHLVACG 186
DB 128 GTLSRYFSRRRLVAVGLALTGNAGSSLLAPALQFLDTFGWRGALLLGAVTLLTPCG 187
OY 187 ALLRPPSLAEDPAVGGPRAQLTS---LLHGGPFLRYVALTLINTGYFIPLYLVAHLQ 242
DB 188 ALLRPLALSGBP-LAPPRTPLAALGLGLFKRAFSVFALGTALIGGYFVPVHLGPHAL 246
OY 243 DLDWDPLPAFLLSVVAISDLVGRVSGWLGDAVPGVTRLMLMTTLTGVSALFPVAQ 302
DB 247 DQGMGYGAALVVAVAAGDACARLASGWLADQGWVPLPRLLVFGSLTG--LGVLAMGL 304
OY 303 APTA-----LVALAFAVGFSGALAPLAFSVLPBELIGTRRIYCGGLQMIETIGG 353
DB 305 VPTVGTGEGWGAFLLAAGAYGLSAGSYAPLVFGLVGLVIGGVQATGLVMMLSLGG 364
OY 354 LUGPPLSGYLRDVSGNVTASFVAVAGAFLLSGSGLTLTPHFFCFSTTSGPQDLVTEALD 413
DB 365 LUGPPLSGFLRDKTGDPSASFLVCSSFILSGSFITYMGLPRALPSCRASPAPPPERGE 424
OY 414 TKVPLPK EGL--EGGLNS 429
DB 425 L-LPVPQVSLSAGGTGS 441

RESULT 9
DB NCCK7_HUMAN PRELIMINARY; PRT; 471 AA.
OY NCCK7; 01-OCT-2002, integrated into UniProtKB/TrEMBL.
DB 01-OCT-2002, sequence version 1.
OY 07-MAR-2006, entry version 21.
DB Hypothetical protein FLJ90193 (Soluble carrier family 16, member 11).
OY Name=SLC16A11;
DB Homo sapiens (Human).
OY Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DB Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OY Homo; NCBI_TaxID=9606;
DB NCBI_TaxID=9606;
P NUCLEOTIDE SEQUENCE.
C TISSUE=Mammary gland;
X PubMed=16303743; DOI=10.1093/dnares/12.2.117;
A Otsuki T., Ota T., Nishikawa T., Hayashi K., Suzuki Y., Yamamoto J.,
A Wakamatsu A., Kimura K., Sakamoto K., Hatano N., Kawai Y., Ishii S.,
A Saito K., Kojima S., Sugiyama T., Ono T., Okano K., Yoshikawa Y.,
A Aotsuka S., Sasaki N., Hattori A., Okumura K., Nagai K., Sugano S.,
A Isogai T.;
T "Signal sequence and keyword trap in silico for selection of full-length human cDNAs encoding secretion or membrane proteins from oligo-

RT capped cDNA libraries."; DNA Res. 12:117-126 (2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung and heart;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RL [3]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Lung and heart;
RC NIH MGC Project;
RG Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
RL
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC
CC EMBL; AK074674; BAC1128.1; -; mRNA.
DB EMBL; BC093886; AAH93886.1; -; mRNA.
DB EMBL; BC093860; AAH93860.1; -; mRNA.
DR Ensemble; ENSG00000174326; Homo sapiens.
DR HGNC; HGNC:23093; SLC16A11.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS50850; MFS; 1.
KW Membrane; Transmembrane; Transport.
SQ SEQUENCE 471 AA; 47791 MW; C586C4FE7A00B626 CRC64;

Query Match 39.9%; Score 930.5; DB 2; Length 471;
Best Local Similarity 45.9%; Pred. No. 3.7e-57;
Matches 206; Conservative 67; Mismatches 145; Indels 31; Gaps 8;

OY 7 PPDGGRVVVLSAFQSAIVFGLRSFGVFVEFVAAFEEOAARVSWIASIGIAVQOFG 66
DB 32 PPDGGRVVVLSAFVAVNGLSYGLRLSLGLALPDLAHFERSAQDTAWVSALALAVQOAA 91
OY 67 SPVGSALSTKFGPRPVMTGILALGMLLASFATSLTHLYLSIGLSSGWSALTFAPTL 126
DB 92 SPVGSALSTRWGARPVMVGVLTSLGLVSAFARSLHLHYLGLGLAGSGWALVFAPAL 151
OY 127 ACLSCYFSRRSLATGLALTGVGLSFTFAPDFQWLSHYAWRGSLLVASLHLVACG 186
DB 152 GTLSRYFSRRRLVAVGLALTGNAGSSLLAPALQFLDTFGWRGALLLGAITVHLTPCG 211
OY 187 ALLRPPSLAEDPAVGGPRAQLT---SLHGGPFLRYVALTLINTGYFIPLYLVAHLQ 242
DB 212 ALLPLVLPGBP-APRSPALALGLSLFTRRAFSIFALGTALVGGGYFVPHLAPHAL 270
OY 243 DLDWDPLPAFLLSVVAISDLVGRVSGWLGDAVPGVTRLMLMTTLTGVSALFPVAQ 299

271 DRGLGCGALVVAAMGADAGARLVCGMLADQGWPVLPRLLAVERGALTGLWVGLVP 330
300 VAQAPTA---LVALAVAYGFTSGALAPLAFSVLPGLIGTRRIYCGILQMIESIGLL 355
331 VVGGEESWGSPLLAAVAYGLSAGSYAPLVFGVLPGLVGVGVQATGVLMMLSIGLL 390
356 GPPLSGYLKRDVSGNYTASFVAVAGAFLLSGGILLTLPHFCCSTTTSGPQDLVTEALDTK 415
391 GPPLSGFLRDEGTGDTASFLLSGSLILSGSFYIGLPR---ALPSCGP-----ASPPA 440
416 VPLPKLEGGLNSTESGPESQS-LTAPG 443
441 TPPPETG-----ELLPAPOAVLLSPG 461
ESULT 10
4KM10_BRARE PRELIMINARY; PRT; 433 AA.
02-AUG-2005, integrated into UniProtKB/TrEMBL.
02-AUG-2005, sequence version 1.
21-FEB-2006, entry version 8.
Hypothetical protein zgc:114041.
ORFNames=zgc:114041;
Brachydanio rerio (Zebrafish) (Danio rerio).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
Cyprinidae; Danio.
NCBI_TaxId=7955;
[1]
NUCLEOTIDE SEQUENCE.
TISSUE=Ovary;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
Rodriguez A.C., Grimwood J., Schmutz J., Skalska U., Smailus D.E.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
NUCLEOTIDE SEQUENCE.
TISSUE=Ovary;
NTH MGC Project;
Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein (By
similarity).
Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
Distributed under the Creative Commons Attribution-NoDerivs license
EMBL; BC098892; AAH98892.1; -, mRNA.
ZFIN; ZDB-GENE-050706-122; zgc:114041.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0016020; C:membrane; IEA.
GO; GO:0005215; F:transporter activity; IEA.
GO; GO:0006810; P:transport; IEA.
InterPro; IPR007114; MFS.
Pfam; PF07690; MFS_1; 1.
PROSITE; PS50850; MFS; 1.

Hypothetical protein; Membrane; Transmembrane; Transport.
SEQUENCE 433 AA; 47067 MW; 364226AE821F7BED CRC64;
Query Match 29.5%; Score 688; DB 2; Length 433;
Best Local Similarity 36.6%; Pred. No. 4,1e-40;
Matches 160; Conservative 76; Mismatches 161; Indels 40; Gaps 6;
QY 6 EPPDGMGRVNVLSAFQSAALVFGLRSFGVFVEFVAAFEQAAARVSWIASIGIAVQOF 65
7 DPPDGGYGVNVVASAFIFIMGLTAVALKNFGLEFLELQNYXSVLTSTSTSLMTSTIAVFHL 66
QY 66 GSPVGSALSTKFGPRPVMTGCGILALCMILASPATSLTHLYSLIGLSGGWALTFAPT 125
67 GSPLASALSMHLSQRPVIMVGGLLAASGMIIASLGSLPMMYLSGVIGLGVSFTWVPA 126
QY 126 LACISCFYSRRSLATGALTGVGLSFTFAPFQWLLSHYAWRGSLLVSAISLHVAC 185
127 NSMVNHYFKRWRIACAIISSGECVFGMAFSPFQWLIESYSWRGALLVIGLIQNLIVC 186
QY 186 GALLRP--PSLAEDPAV-----GGPRAQLT---SLHGGPFLRYTVALTLINTGYFIPY 234
187 GALKMKPLQPVQTSRKAVLDSKEGTGTSKVTFOCSLIQRPPELLYIVFAIPAAAGFFIP 246
QY 235 LHLVAHLQDLWDPLPAFLLSVVAISDLVGRVVSGLGDAVPGVTRLLMLWTLLTGV 294
247 LFLVYVNNLGMDOYMAASLVSLSLADLLGRGCGMLANLCLRLNQLLTWVATAMGV 306
QY 295 LALFPVAQAPTALVALAVAYGFTSGALAPLAFSVLPGLIGTRRIYCGILQMIESIGL 354
307 LLLPIAYGYWSVLVFTSLYGFLEGCVAIHVTSIVDIVGLEGFDLSALGLFMILRSIGCF 366
QY 355 LGPPLSGYLKRDVSGNYTASFVAVAG-AFLLSGGILL-----TLPHFFCF 398
367 LGPPAAGWLVDWTHNFGAIFYLSGICLLLSGCFVVLVDWLVEKKKKKLOETL----- 419
QY 399 TTTSGPQDLVTEALDTK 415
Db 420 -----QDPCTETTDLK 430
RESULT 11
Q801R0 XENLA PRELIMINARY; PRT; 533 AA.
Q801R0;
01-JUN-2003, integrated into UniProtKB/TrEMBL.
01-JUN-2003, sequence version 1.
21-FEB-2006, entry version 19.
Slc16a6-prov protein.
Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
NCBI_TaxId=8355;
[1]
NUCLEOTIDE SEQUENCE.
TISSUE=Embryo;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

TOPO_DOM 315 318 Extracellular (Potential).
 TRANSMEM 319 339 Potential.
 TOPO_DOM 340 352 Cytoplasmic (Potential).
 TRANSMEM 353 373 Potential.
 TOPO_DOM 374 386 Extracellular (Potential).
 TRANSMEM 387 407 Potential.
 TOPO_DOM 408 492 Cytoplasmic (Potential).
 SEQUENCE 492 AA; 51590 MW; 5CCCA4DC38B35FA0 CRC64;

Query Match 27.7%; Score 645; DB 1; Length 492;
 Best Local Similarity 36.1%; Pred. No. 5e-37;
 Matches 170; Conservative 70; Mismatches 193; Indels 38; Gaps 9;

3 RRTPEPPDGGWGRVVVLSAFQOSALVFGVLSFGVFPVEFVAAFEEQAAVSWIASIGIAV 62
 8 RGAGPPDGGWGVVVLGACFVITGFAYGFPKAVSVFFRELKRDGAGYSDTAWVSSIMLAM 67
 63 QQFGSPVGSALSTKFGPRPVVMTGGILALGMLLASFATSLTHLYLSIGLLSGSGWALT 122
 68 LYGTGPLSSILVTRFGCRPVMLAGLLASAGMILASPASRLLELYTAGVLTGLALNF 127
 123 APTLACLSCYFSRRSLATGLALTGVGLSFTFAFPFQWMLSHYAWRGSLLVLSLSHL 182
 128 QPSLIMLGLYFERRRPLANGLAAGSPVFLSTLSPLQLLGERFGWRGFFLLFGLLHC 187
 183 VACGALLRPPLAE---DPAVGPRAOQLTSLH----HGFLRYTVALTLINTGYFI 234
 188 CACGAVMRPPGPQRPDPAPRPGRARHRLDLAVCTDRTFMVYVTKFIMALGLFVPA 247
 235 LHLVAHLQDLWDPLPAFLLSVAISDLGVRVSG--WLGDAVPPVTRLLMLWTTLT 291
 248 ILLVNYAKDAGVPDAEAAFLLSIVGFVDIVARPACGALAGLRLP-HVPYLFSLALLAN 306
 292 GVSLLAFVPAQPTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGILQMIESI 351
 307 GLTDLISARARSYGLTVAFCIAFGLSYGMWALQFEVLMATVGAPRFPALGLVLLVEAV 366
 352 GGLGPPLSGYLKRVSGNYTASFVVAGAFLLSGSGILLTLPHFCFSTT----- 400
 367 AVLIGPPSAGRLVDALKNYEIIFYLAGS-EVALAGVFMAVTTYCCLRCSKNISSGRSAEG 425
 401 -TSGPQDLVTEALDTKVPLPKEGLEGGLNSTSGPESQSLTAPGLLPRLG 450
 426 GASDPEDV--EAERDSEMPA-----STE---EPGSLALEVLSPRAG 463

:SULT 13
 :JJC0 MOUSE PRELIMINARY; PRT; 290 AA.
 :Q9JJC0;
 :01-OCT-2000, integrated into UniProtKB/TrEMBL.
 :01-OCT-2000, sequence version 1.
 :07-FEB-2006, entry version 21.
 :Mus musculus brain cDNA, clone MNcb-2717 (Novel RNA polymerase Rpb1 C-
 :terminal repeat domain containing protein).
 :Name=Slc16a11; Synonyms=RP23-198E14.4; ORFNames=RP23-198E14.4-003;
 :Mus musculus (Mouse).
 :Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 :Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 :Muridea; Muridae; Murinae; Mus.
 :NCBI_TaxID=10090;
 :[1]

:NUCLEOTIDE SEQUENCE.
 :STRAIN=C57BL;
 :Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
 :Hashimoto K.;
 :Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 :[2]
 :NUCLEOTIDE SEQUENCE.
 :Sycamore N.;
 :Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 :-----
 :Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms

CC Distributed under the Creative Commons Attribution-NoDerivs license

 CC EMBL; AB041591; BAA95074.1; -; mRNA.
 DR EMBL; AL669869; CA135221.1; -; Genomic DNA.
 DR Ensembl; ENSMUSG0000040938; Mus musculus.
 DR MGI; MGI:2663709; Slc16a11.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR011701; MFS_1.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PS50850; MFS; 1.
 KW Membrane; Transmembrane; Transport.
 SQ SEQUENCE 290 AA; 29031 MW; AA5501FCA7ED7F9F CRC64;

Query Match 27.6%; Score 643; DB 2; Length 290;
 Best Local Similarity 47.4%; Pred. No. 3.9e-37;
 Matches 138; Conservative 43; Mismatches 94; Indels 16; Gaps 4;

QY 84 MTGILALGMLLASFATSLTHLYLSIGLLSGSGMALTFAPTACLSCYFSRRSLATGL 143
 DB 1 MVGGVLTSLGLVFSAFARSLHLHYLGLGLLAGSGMALVFAPALGTLSTRYFSRRVLA 60
 QY 144 ALTGVGLSFTFAFPFQWMLSHYAWRGSLLVLSLSHLVACGALLRPPSLAEDPAVGP 203
 DB 61 ALTNGASSLLAPALQFLDFTFGWRGALLLGAVTIHLTPCGALLRPLALSGDP-LADP 119
 QY 204 RAQLTS---LHNGFLRYTVALTLINTGYFIPLYHLVAHLQDLWDPLPAFLLSVVA 259
 DB 120 RTPLAALGLGLFKRAFSVFALGTALIGGYFVPYVHLGPALDQGMGYGAALVVAVA 179
 QY 260 ISDLGVRVSGMLGDVAPGPVTRLLMLWTTLTGVSLLAFVPAQPTA-----LVAL 310
 DB 180 VGDACARLASGWLADQGWPLRLLLVFGSLTG-LGVLAMGLVPTVGTBEGWGA 237
 QY 311 AVAYGFTSGALAPLAFSVLPBELIGTRRIYCGILQMIESIGLLGPPLSG 361
 DB 238 AGAYGLSAGSYAPLVFGLVPLGLVIGGVQATGLVNMMLSLGGLGPPLSG 288

RESULT 14
 Q8BGC3 MOUSE PRELIMINARY; PRT; 486 AA.
 AC Q8BGC3;
 DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
 DT 01-MAR-2003, sequence version 1.
 DT 07-FEB-2006, entry version 21.
 DE 14, 17 days embryo head cDNA, RIKEN full-length enriched library,
 DE clone:3230401A21 product:hypothetical protein, full insert sequence
 DE (10 days neonate skin cDNA, RIKEN full-length enriched library,
 DE clone:4732488P05 product:hypothetical protein, full insert sequence).
 GN Name=Slc16a12;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head, and Skin;
 RX PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,

RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Acuraliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnieceki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasaki Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Sempke C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yaagi K.,
RA Yamaniishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Bruscia V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Ohno Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.,
RA "The transcriptional landscape of the mammalian genome.",
RA Science 309:1559-1563(2005).
RA [3]

NUCLEOTIDE SEQUENCE.
RA STRAIN=C57BL/6J; TISSUE=Head, and Skin;
RA PubMed=16141073; DOI=10.1126/science.1112009;
RA RIKEN Genome Exploration Research Group, and Genome Science Group
RA (Genome Network Core Team) and the FANTOM Consortium;
RA "Antisense Transcription in the Mammalian Transcriptome";
RA Science 309:1564-1566(2005).
RA [4]

NUCLEOTIDE SEQUENCE.
RA STRAIN=C57BL/6J; TISSUE=Head, and Skin;
RA MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,
RA Yaagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasaki Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirokawa K., Kawai K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN [5]

NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head, and Skin;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [6]

NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head, and Skin;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes";
RL Genome Res. 10:1617-1630(2000).
RN [7]

NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head, and Skin;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer";
RL Genome Res. 10:1757-1771(2000).
RN [8]

NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head, and Skin;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numata K., Nishizaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NonCommercial license
EMBL; AK029076; BAC26279.1; -; mRNA.
DR EMBL; AK028284; BAC25857.1; -; mRNA.
DR Ensembl; ENSMUSG0000009378; Mus musculus.

MGI; MGI:2147716; S1c16a12.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0016020; C:membrane; IEA.
GO; GO:0005215; F:transporter activity; IEA.
GO; GO:0006810; P:transport; IEA.
InterPro; IPR007114; MFS.
InterPro; IPR011701; MFS_1.
Pfam; PF07690; MFS_1; 1.
PROSITE; PS50850; MFS; 1.
Hypothetical protein; Membrane; Transmembrane; Transport.
SEQUENCE 486 AA; 53152 MW; AA3CC9951FB28247 CRC64;

Very Match 27.5%; Score 640; DB 2; Length 486;
Best Local Similarity 32.1%; Pred. No. 1.1e-36;
Matches 152; Conservative 89; Mismatches 172; Indels 60; Gaps 7;

7 PPDGGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEQAAARVSWIASIGIAVQOFG 66
12 PPDGGRVMMIVAGCFLVTICTRAVTRCISIFVEFQTYFAQDYSQTAWIHSIVDCMTMLC 71
67 SPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSGWAFTAPTL 126
72 APLGSVVSNOQSCQAGIMLGGLLASTGFIKLSFATSLKHLVLSLGLVTLGFGALCYSPAI 131
127 ACLSCYFSRRRSIATGLALTGVGLSFTFAPFPQWLLSHYAWRGSLLVSALSLHVACG 186
132 AMVKYFSRRKALAYGIAMSGSIGTIFLAPVVQLLIEQFSWRGALLIGGFVLNLCVCG 191
187 ALRPPSLAEDPAV-----GGPRAQLT-----SLHGGPFLRYT 220
192 ALMRPITLKEDRSVDEKHNHRESQREDCKQASPYSPLTKECTETRLCCSLQOEYGFLLMS 251
221 VALTLINTGYFIPIY-----LHLVAHLQDLWDPLPAFLLSVVAISDLVGRVVSGLGD 274
252 DFVVLAVSVLFMAYGCSPLFVYLVPYALSVGSHQAFLMSILGVIDIVGNITFGWLTD 311
275 AVPGPVTRL-----MLWTLTGVSLLAFVQAAPTALVALAVAGFTGALAPLAF 326
312 -----RRCLKNRYVCYLFVALDGLCYLCLPMLQTFPLLVPSCTFGYFDGAVVTLIP 365
327 SVLPGLICTRRIYCGLLQMIIESIGLGPPLSGYLKRDVSGNYTASFVVAAGFLLSGSG 386
366 VVTAEIVGTSLSSALGVVYFLHAVPYLVSPPIAGMLVDVTGSYTAFFLLCG-FAMIFSS 424
387 ILLTLPHFPCFSTTSGFODLVTEALDTKVPPLPKEGLEGLNSTESGPESQSL 439
425 ILL-----GPRVIVKRMKRTQVPFVKSDPKLQLMTNGSVAYSV 464

SULT 15
Q5T7J2 HUMAN PRELIMINARY; PRT; 463 AA.
Q5T7J2;

21-DEC-2004, integrated into UniProtKB/TrEMBL.
21-DEC-2004, sequence version 1.
07-FEB-2006, entry version 11.
Solute carrier family 16 (Monocarboxylic acid transporters), member 12 (Fragment).
ORFNames=RP11-168010.10-001;
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
Homo.
NCBI_TaxID=9606;

NUCLEOTIDE SEQUENCE.
Chapman J.;
Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.

Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
Distributed under the Creative Commons Attribution-NonDerivs License
EMBL; AL353146; CA112383.1; -; Genomic_DNA.

DR Ensembl; ENSG00000152779; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS50850; MFS; 1.
FT NON TER 463 463
SQ SEQUENCE 463 AA; 50570 MW; C38CA95E85D338DF CRC64;

Query Match 27.3%; Score 637.5; DB 2; Length 463;
Best Local Similarity 33.5%; Pred. No. 1.6e-36;
Matches 146; Conservative 83; Mismatches 144; Indels 63; Gaps 9;

QY 2 ARTEPPDGGWRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEQAAARVSWIASIGIA 61
Db 37 ARSTSPDGGWMMIVAGCFLVTICTRAVTRCISIFVEFQTYFTQDYAQTAWIHSIVDC 96
QY 62 VQGFSPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSGWAFT 121
Db 97 VTMLCAPLGSVSNHLSQVGYIMLGGLLASTGLISSFATSLKHLVLSLGLVTLGFGALC 156
QY 122 FAPTLACLSCYFSRRRSIATGLALTGVGLSFTFAPFPQWLLSHYAWRGSLLVSALSLH 181
Db 157 YSPALAMVKYFSRRKALAYGIAMSGSIGTIFLAPVVQLLIEQFSWRGALLIGGFVLN 216
QY 182 LVACGALLRPPSLAEDPAVG-----GGPRAQLT-----S 209
Db 217 LCVCGALMRPITLKEDHTTPEQNHVCRCTOKEDIKRVSPYSSLTKEWAQTCCLCCLOEYS 276
QY 210 LHLGGPFLRYTVALTLINTG-----YFIPI-LHL-VAHLQDLWDPLPAFLLSVVAIS 261
Db 277 FLMSDFVVLAVSVLFMAYGCSPLFVYLVPYALSVGSHQ-----AAFLMSILGVI 328
QY 262 DLVGRVVSGLGDVAVPGPVTRL-----MLWTLTGVSLLAFVQAAPTALVALAVA 313
Db 329 DIIGNTFPGWLTD-----RRCLKNYQVVCYLFVAGMDGLCYLCLPMLQSLPLLVPSCT 382
QY 314 YGFTSGALAPLAFSVLPGLICTRRIYCGLLQMIIESIGLGPPLSGYLKRDVSGNYTAS 373
Db 383 FGYPDGAIVTLLPVVTEIVGTSLSSALGVVYFLHAVPYLVSPPIAGRLVDVTGSYTAA 442
QY 374 FVVAAGFLLSGSGLL 389
Db 443 FLLCG-FSMIFSSVLL 457

Search completed: September 7, 2006, 13:07:17
Job time : 305 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

3M protein - protein search, using sw model

Run on: September 7, 2006, 13:08:16 ; Search time 182 Seconds
(without alignments)
1163.127 Million cell updates/sec

Title: US-10-654-428-2
Perfect score: 2331
Sequence: 1 MARRTEPPDGGWGRVVVLSA.....SLTAPGLLPRLGLHRTTVP 457
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2331	100.0	457	US-09-829-432-2	Sequence 2, Appli
2	2331	100.0	457	US-10-654-428-2	Sequence 2, Appli
3	2154	92.4	426	US-10-451-168-110	Sequence 110, App
4	2154	92.4	426	US-10-980-387-110	Sequence 110, App
5	2151	92.3	426	US-10-297-022-16	Sequence 16, Appli
6	2150	92.2	426	US-10-072-012-306	Sequence 306, App
7	2142.5	91.9	427	US-10-072-012-338	Sequence 338, App
8	2142.5	91.9	427	US-10-072-012-779	Sequence 779, App
9	2142	91.9	426	US-10-072-012-302	Sequence 302, App
10	2142	91.9	426	US-10-072-012-304	Sequence 304, App
11	930.5	39.9	447	US-10-072-012-24	Sequence 24, Appli
12	920	39.5	472	US-10-072-012-30	Sequence 30, Appli
13	704	30.2	142	US-10-644-807-224	Sequence 224, App
14	704	30.2	142	US-10-644-807-314	Sequence 314, App
15	645	27.7	456	US-09-829-432-4	Sequence 4, Appli
16	645	27.7	456	US-10-654-428-4	Sequence 4, Appli
17	643	27.6	290	US-10-072-012-1391	Sequence 1391, Ap
18	637.5	27.3	486	US-10-419-190-2	Sequence 2, Appli
19	637.5	27.3	516	US-10-380-727-12	Sequence 12, Appli
20	637.5	27.3	516	US-10-511-538-87	Sequence 87, Appli
21	631	27.1	456	US-09-829-432-5	Sequence 5, Appli
22	631	27.1	456	US-10-654-428-5	Sequence 5, Appli
23	630	27.0	308	US-10-343-903-1	Sequence 1, Appli
24	625.5	26.8	504	US-10-272-196-3	Sequence 3, Appli
25	625.5	26.8	513	US-10-272-196-39	Sequence 39, Appli
26	617	26.5	489	US-10-072-012-339	Sequence 339, App
27	617	26.5	489	US-10-072-012-340	Sequence 340, App

28	617	26.5	489	4	US-10-072-012-782	Sequence 782, App
29	617	26.5	489	4	US-10-072-012-783	Sequence 783, App
30	609	26.1	484	4	US-10-072-012-341	Sequence 341, App
31	608	26.1	478	4	US-10-072-012-781	Sequence 781, App
32	608	26.1	478	5	US-10-723-860-1841	Sequence 1841, Ap
33	608	26.1	478	5	US-10-756-149-5185	Sequence 5185, Ap
34	605.5	26.0	478	4	US-10-072-012-780	Sequence 780, App
35	605.5	26.0	478	4	US-10-272-196-2	Sequence 2, Appli
36	605.5	26.0	487	4	US-10-272-196-38	Sequence 38, Appli
37	600.5	25.8	197	4	US-10-363-829-321	Sequence 321, App
38	600.5	25.8	523	4	US-10-295-027-70	Sequence 70, Appli
39	600.5	25.8	523	4	US-10-058-270A-42	Sequence 42, Appli
40	600.5	25.8	562	5	US-10-128-558-225	Sequence 225, App
41	599.5	25.7	523	4	US-10-241-220-95	Sequence 95, Appli
42	599.5	25.7	523	4	US-10-295-027-72	Sequence 72, Appli
43	599.5	25.7	523	5	US-10-872-972-95	Sequence 95, Appli
44	599.5	25.7	523	5	US-10-872-991-95	Sequence 95, Appli
45	599.5	25.7	523	5	US-10-756-149-5146	Sequence 5146, Ap

ALIGNMENTS

RESULT 1
US-09-829-432-2
Sequence 2, Application US/09829432
General Information:
APPLICANT: KETCHUM, Karen et al.
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: CL001013
CURRENT APPLICATION NUMBER: US/09/829,432
CURRENT FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 09/815,301
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: 60/254,554
PRIOR FILING DATE: 2000-12-12
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 457
TYPE: PRT
ORGANISM: Human
US-09-829-432-2

Query Match 100.0%; Score 2331; DB 3; Length 457;
Best Local Similarity 100.0%; Pred. No. 3.8e-194;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MARRTEPPDGGWGRVVVLSAFFQSALVFGVIRSFVFEVFAAFEEQAAARVSWIASIGI	60
DB	1	MARRTEPPDGGWGRVVVLSAFFQSALVFGVIRSFVFEVFAAFEEQAAARVSWIASIGI	60
QY	61	AVOQFGSPVGSALSTKFGPRPVMTGGILAAIGMLASFATSLTHLYLSIGLSSGMAI	120
DB	61	AVOQFGSPVGSALSTKFGPRPVMTGGILAAIGMLASFATSLTHLYLSIGLSSGMAI	120
QY	121	TFAPTLACSCYFSRRSLATGLALTGVGLSFTFAPFFQWLLSHYAWRGSLIVSALSL	180
DB	121	TFAPTLACSCYFSRRSLATGLALTGVGLSFTFAPFFQWLLSHYAWRGSLIVSALSL	180
QY	181	HLVACGALLRPPSLAEDPAVGPRQQLTSLHHGPFRLRYTVALTLINTGYFFPYLHLVAH	240
DB	181	HLVACGALLRPPSLAEDPAVGPRQQLTSLHHGPFRLRYTVALTLINTGYFFPYLHLVAH	240
QY	241	LQDLMDPLPAAPFLSVVAISDLVGRVSGWLGDAVPGPVTRLMLWTTLTGVSALFPV	300
DB	241	LQDLMDPLPAAPFLSVVAISDLVGRVSGWLGDAVPGPVTRLMLWTTLTGVSALFPV	300
QY	301	AQAPTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGLGILQMIISIGLGPPLS	360
DB	301	AQAPTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGLGILQMIISIGLGPPLS	360

b 301 AQAFTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGLGLLQMIESIGGLGPPLS 360
y 361 GYLRDVSGNYTASFVVAGAFLLSGSGILLTLPHFCEFSTTTSGPQDLVTEALDTKVPLPK 420
b 361 GYLRDVSGNYTASFVVAGAFLLSGSGILLTLPHFCEFSTTTSGPQDLVTEALDTKVPLPK 420
y 421 EGLEGGLNSTESGPESQSLTAPGLLPRLGLHRTTVP 457
b 421 EGLEGGLNSTESGPESQSLTAPGLLPRLGLHRTTVP 457

ESULT 2
S-10-654-428-2

Sequence 2, Application US/10654428
Publication No. US20040082035A1
GENERAL INFORMATION:

APPLICANT: KETCHUM, Karen et al.
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: CL001013CIP-CON
CURRENT APPLICATION NUMBER: US/10/654,428
CURRENT FILING DATE: 2003-09-04
PRIOR APPLICATION NUMBER: 09/829,432
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 09/815,301
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: 60/254,554
PRIOR FILING DATE: 2000-12-12
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 2
LENGTH: 457
TYPE: PRT
ORGANISM: Homo sapiens
S-10-654-428-2

Query Match 100.0%; Score 2331; DB 4; Length 457;
Best Local Similarity 100.0%; Pred. No. 3.8e-194;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLSFGVFEFVAAFEEOAARVSWIASIGI 60
b 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLSFGVFEFVAAFEEOAARVSWIASIGI 60
y 61 AVQQFGSPVGSALSTKFGPRPVVMTGGILALGMLLASFATSLTHLYSLIGLSSGSWAL 120
b 61 AVQQFGSPVGSALSTKFGPRPVVMTGGILALGMLLASFATSLTHLYSLIGLSSGSWAL 120
y 121 TFAPTLACLSCYFSRRSLATGALTGVGSSFTFAPFQWLLSHYAMRGSLLLVSALS 180
b 121 TFAPTLACLSCYFSRRSLATGALTGVGSSFTFAPFQWLLSHYAMRGSLLLVSALS 180
y 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFLRYTVALTLINTGYFIPYLHLVAH 240
b 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFLRYTVALTLINTGYFIPYLHLVAH 240
y 241 LQDLWDPLPAAFLLSVVAISDLVGRVVGWLGDAVPGPVTRLMLMTTLTGVSALFPV 300
b 241 LQDLWDPLPAAFLLSVVAISDLVGRVVGWLGDAVPGPVTRLMLMTTLTGVSALFPV 300
y 301 AQAFTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGLGLLQMIESIGGLGPPLS 360
b 301 AQAFTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGLGLLQMIESIGGLGPPLS 360
y 361 GYLRDVSGNYTASFVVAGAFLLSGSGILLTLPHFCEFSTTTSGPQDLVTEALDTKVPLPK 420
b 361 GYLRDVSGNYTASFVVAGAFLLSGSGILLTLPHFCEFSTTTSGPQDLVTEALDTKVPLPK 420
y 421 EGLEGGLNSTESGPESQSLTAPGLLPRLGLHRTTVP 457
b 421 EGLEGGLNSTESGPESQSLTAPGLLPRLGLHRTTVP 457

RESULT 3
US-10-451-168-110
; Sequence 110, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM p.l.c.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 110
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-451-168-110

Query Match 92.4%; Score 2154; DB 4; Length 426;
Best Local Similarity 99.8%; Pred. No. 8.9e-179;
Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLSFGVFEFVAAFEEOAARVSWIASIGI 60
Db 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLSFGVFEFVAAFEEOAARVSWIASIGI 60
QY 61 AVQQFGSPVGSALSTKFGPRPVVMTGGILALGMLLASFATSLTHLYSLIGLSSGSWAL 120
Db 61 AVQQFGSPVGSALSTKFGPRPVVMTGGILALGMLLASFATSLTHLYSLIGLSSGSWAL 120
QY 121 TFAPTLACLSCYFSRRSLATGALTGVGSSFTFAPFQWLLSHYAMRGSLLLVSALS 180
Db 121 TFAPTLACLSCYFSRRSLATGALTGVGSSFTFAPFQWLLSHYAMRGSLLLVSALS 180
QY 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFLRYTVALTLINTGYFIPYLHLVAH 240
Db 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFLRYTVALTLINTGYFIPYLHLVAH 240
QY 241 LQDLWDPLPAAFLLSVVAISDLVGRVVGWLGDAVPGPVTRLMLMTTLTGVSALFPV 300
Db 241 LQDLWDPLPAAFLLSVVAISDLVGRVVGWLGDAVPGPVTRLMLMTTLTGVSALFPV 300
QY 301 AQAFTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGLGLLQMIESIGGLGPPLS 360
Db 301 AQAFTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGLGLLQMIESIGGLGPPLS 360
QY 361 GYLRDVSGNYTASFVVAGAFLLSGSGILLTLPHFCEFSTTTSGPQDLVTEALDTKVPLPK 420
Db 361 GYLRDVSGNYTASFVVAGAFLLSGSGILLTLPHFCEFSTTTSGPQDLVTEALDTKVPLPK 420
QY 421 EGLE 424
Db 421 EGLE 424

Db 421 EGLE 424

RESULT 4

US-10-980-387-110
; Sequence 110, Application US/10980387
; Publication No. US20050164343A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM p.l.c.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/980,387
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US/10/980,387
; PRIOR FILING DATE: 2003-11-0312
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; JS-10-980-387-110

Query Match 92.4%; Score 2154; DB 5; Length 426;
Best Local Similarity 99.8%; Pred. No. 8.9e-179;
Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLSFGVFEVFEVAAFEEOAARVSWIASIGI 60
1 MARRTEPPDGGWGVVVLSAFFQSALVFGVLSFGVFEVFEVAAFEEOAARVSWIASIGI 60
61 AVOQFGSPVGSALSTKFGPRPVMTGGILAAAGMLLASFATSLTHLYSLIGLLSGSGMAL 120
61 AVOQFGSPVGSALSTKFGPRPVMTGGILAAAGMLLASFATSLTHLYSLIGLLSGSGMAL 120
61 AVOQFGSPVGSALSTKFGPRPVMTGGILAAAGMLLASFATSLTHLYSLIGLLSGSGMAL 120
121 TFAPTLACLSCTFSRRSLATGTLATGVLSSFTFAPFFQWLLSHYAMRGSLLVSALSL 180
121 TFAPTLACLSCTFSRRSLATGTLATGVLSSFTFAPFFQWLLSHYAMRGSLLVSALSL 180
121 TFAPTLACLSCTFSRRSLATGTLATGVLSSFTFAPFFQWLLSHYAMRGSLLVSALSL 180
181 HLVACGALLRPPSLAEDPAVGPRAOQLTSLHGPFLRYTVALTLINTGYFIPYLHVAH 240
181 HLVACGALLRPPSLAEDPAVGPRAOQLTSLHGPFLRYTVALTLINTGYFIPYLHVAH 240
181 HLVACGALLRPPSLAEDPAVGPRAOQLTSLHGPFLRYTVALTLINTGYFIPYLHVAH 240
241 LODLDWDLPAFAFLLSVAISDLVGRVSGWLGDAVPGPVTRLMLMTLLTGVSLALFPV 300
241 LODLDWDLPAFAFLLSVAISDLVGRVSGWLGDAVPGPVTRLMLMTLLTGVSLALFPV 300
241 LODLDWDLPAFAFLLSVAISDLVGRVSGWLGDAVPGPVTRLMLMTLLTGVSLALFPV 300
301 AQAPTALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGILQMIIESIGLLGPPLS 360
301 AQAPTALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGILQMIIESIGLLGPPLS 360
301 AQAPTALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGILQMIIESIGLLGPPLS 360
361 GYLKDVSGNYTASFVAVAGAFLLSGGILLTLPHFFCFSTTTSGPQDLVTEALDTKVP LK 420
|||||

Db 361 GYLKDVSGNYTASFVAVAGAFLLSGGILLTLPHFFCFSTTTSGPQDLVTEALDTKVP LK 420

QY 421 EGLE 424

Db 421 EGLE 424

RESULT 5

US-10-297-022-16
; Sequence 16, Application US/10297022
; Publication No. US20030216310A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: THORNTON, Michael
; APPLICANT: WALIA, Narinder K.
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Danniel B.
; APPLICANT: LAL, Preeti
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: TANG, Y. Tom
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BRUNS, Christopher M.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: YANG, Junming
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: SANJANWALA, Madhu S.
; APPLICANT: RAUMANN, Brigitte E.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: GREENE, Barrie D.
; APPLICANT: KEARNEY, Liam
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: SEILHAMER, Jeffrey J.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: BURFORD, Neil
; APPLICANT: DING, Li
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: HILLMAN, Jennifer L.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0109 PCT
; CURRENT APPLICATION NUMBER: US/10/297,022
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 60/208,424; 60/209,001; 60/210,588; 60/212,335; 60/213,747;
; PRIOR FILING DATE: 2000-05-26; 2000-06-01; 2000-06-08; 2000-06-16; 2000-06-22; 2000-06-28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030216310A1 817477CD1
US-10-297-022-16

Query Match 92.3%; Score 2151; DB 4; Length 426;
Best Local Similarity 99.5%; Pred. No. 1.6e-178;
Matches 422; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLSFGVFEVFEVAAFEEOAARVSWIASIGI 60
1 MARRTEPPDGGWGVVVLSAFFQSALVFGVLSFGVFEVFEVAAFEEOAARVSWIASIGI 60
61 AVOQFGSPVGSALSTKFGPRPVMTGGILAAAGMLLASFATSLTHLYSLIGLLSGSGMAL 120
61 AVOQFGSPVGSALSTKFGPRPVMTGGILAAAGMLLASFATSLTHLYSLIGLLSGSGMAL 120
61 AVOQFGSPVGSALSTKFGPRPVMTGGILAAAGMLLASFATSLTHLYSLIGLLSGSGMAL 120

Y 121 TFAPTLACLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLISHYAWRGSLLSALSL 180
Y 121 TFAPTLACLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLISHYAWRGSLLSALSL 180
Y 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYHLVAH 240
Y 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYHLVAH 240
Y 241 LQDLWDPLPAFAFLLSVVAISDLVGRVVGWLGDVAVPGPVTRLMLMTTLTGVSALFPV 300
Y 241 LQDLWDPLPAFAFLLSVVAISDLVGRVVGWLGDVAVPGPVTRLMLMTTLTGVSALFPV 300
Y 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPBLIGTRRIYCGLGILQMIIESIGLLGPPLS 360
Y 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPBLIGTRRIYCGLGILQMIIESIGLLGPPLS 360
Y 361 GYLKDVSGNYTASFVAVAGAFLLSGSGILLTLPHFCESTTTSGPQDLVTEALDTKVPLPK 420
Y 361 GYLKDVSGNYTASFVAVAGAFLLSGSGILLTLPHFCESTTTSGPQDLVTEALDTKVPLPK 420
Y 421 EGLE 424
Y 421 EGLE 424
Y 421 EGLE 424

ESULT 6
S-10-072-012-306
Sequence 306, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Paturajan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimlr Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057

;; PRIOR FILING DATE: 2001-02-07
;; PRIOR APPLICATION NUMBER: 60/266,975
;; PRIOR FILING DATE: 2001-02-07
;; PRIOR APPLICATION NUMBER: 60/267,459
;; PRIOR FILING DATE: 2001-02-08
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1391
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 306
;; LENGTH: 426
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-072-012-306
Query Match 92.2%; Score 2150; DB 4; Length 426;
Best Local Similarity 99.5%; Pred. No. 2e-178;
Matches 422; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGI 60
Db 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGI 60
QY 61 AVOQFGSPVGSALSTKEGPRPVMTGGILALGMLLSFATSLTHLYLSIGLSSGWAL 120
Db 61 AVOQFGSPVGSALSTKEGPRPVMTGGILALGMLLSFATSLTHLYLSIGLSSGWAL 120
QY 121 TFAPTLACLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLISHYAWRGSLLSALSL 180
Db 121 TFAPTLACLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLISHYAWRGSLLSALSL 180
QY 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYHLVAH 240
Db 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYHLVAH 240
QY 241 LQDLWDPLPAFAFLLSVVAISDLVGRVVGWLGDVAVPGPVTRLMLMTTLTGVSALFPV 300
Db 241 LQDLWDPLPAFAFLLSVVAISDLVGRVVGWLGDVAVPGPVTRLMLMTTLTGVSALFPV 300
QY 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPBLIGTRRIYCGLGILQMIIESIGLLGPPLS 360
Db 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPBLIGTRRIYCGLGILQMIIESIGLLGPPLS 360
QY 361 GYLKDVSGNYTASFVAVAGAFLLSGSGILLTLPHFCESTTTSGPQDLVTEALDTKVPLPK 420
Db 361 GYLKDVSGNYTASFVAVAGAFLLSGSGILLTLPHFCESTTTSGPQDLVTEALDTKVPLPK 420
QY 421 EGLE 424
Db 421 EGLE 424
RESULT 7
US-10-072-012-338
Sequence 338, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Paturajan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimlr Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A

APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 338
LENGTH: 427
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (15)
OTHER INFORMATION: wherein Xaa is any amino acid.
S-10-072-012-338

Query Match 91.9%; Score 2142.5; DB 4; Length 427;
Best Local Similarity 99.3%; Pred. No. 9e-178;
Matches 422; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
1 MARRTEPPDGGWG-RVVVLSAFQSAIVFGLRSFGVFVEFVAAFEEQARVSWIASIG 59
1 MARRTEPPDGGWGXVVVLSAFQSAIVFGLRSFGVFVEFVAAFEEQARVSWIASIG 60
60 IAVQFGSPVGSALSTKFGPRPVMTGILALGMLLASFATSLTHLYLSIGLSSGWA 119
61 IAVQHGSPVGSALSTKFGPRPVMTGILALGMLLASFATSLTHLYLSIGLSSGWA 120
120 LTFAPTLACLSYFSRRSLATGALTGVGLSSFTFAPFQWLLSHYAMRGSLIVSALS 179
121 LTFAPTLACLSYFSRRSLATGALTGVGLSSFTFAPFQWLLSHYAMRGSLIVSALS 180
180 LHLVACGALLRPSLAEDPAVGPRQOLTSLHNGPFLRYTVALTLINTGYFIPYLHLVA 239
181 LHLVACGALLRPSLAEDPAVGPRQOLTSLHNGPFLRYTVALTLINTGYFIPYLHLVA 240
240 HLQDLWDPLPAFLSLVAISDLVGRVSWGWDGAVGPGVTRLLMLMTTLTGVSIALFP 299
241 HLQDLWDPLPAFLSLVAISDLVGRVSWGWDGAVGPGVTRLLMLMTTLTGVSIALFP 300
300 VAOAPTALVALAVAGFTSGALAPLAFSVLPFLIGTRRIYCGILQMTESIGLLGPPL 359
301 VAOAPTALVALAVAGFTSGALAPLAFSVLPFLIGTRRIYCGILQMTESIGLLGPPL 360
360 SGYLRDVSGNYTASFVAGAFLLSGGILLTLPHFCFSTTSGPQDLVTEALDTKVPLP 419

Db 361 SGYLRDVSGNYTASFVAGAFLLSGGILLTLPHFCFSTTSGPQDLVTEALDTKVPLP 420
QY 420 KEGLE 424
Db 421 KEGLE 425

RESULT 8
US-10-072-012-779
Sequence 779, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigar, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 779
LENGTH: 427
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (15)
OTHER INFORMATION: wherein Xaa is any amino acid.
US-10-072-012-779

Query Match 91.9%; Score 2142.5; DB 4; Length 427;

Best Local Similarity 99.3%; Pred. No. 9e-178;
Matches 422; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

1 MARRTEPPDGGWG-RVVVLSAFFQSALVFGVLRSEGVFVEFVAAAFEEQAAARVSWIASIG 59
1 MARRTEPPDGGWGXVVVLSAFFQSALVFGVLRSEGVFVEFVAAAFEEQAAARVSWIASIG 60
60 IAVQFGSPVGSALSTKFGPRPVMTGGILAAAGMLASFATSLTHLYLSIGLSGSWA 119
61 IAVQFGSPVGSALSTKFGPRPVMTGGILAAAGMLASFATSLTHLYLSIGLSGSWA 120
120 LTFAPTLACLSCTFSRRSLATGLALTGVGLSSFTFAPFQWLLSHYAMRGSLLVSALS 179
121 LTFAPTLACLSCTFSRRSLATGLALTGVGLSSFTFAPFQWLLSHYAMRGSLLVSALS 180
180 LHLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYLHVA 239
181 LHLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYLHVA 240
240 HLQDLMDPLPAAFLLSVVAISDLVGRVSWGMDAVPGPVTRLMLMTTLTGVSIALFP 299
241 HLQDLMDPLPAAFLLSVVAISDLVGRVSWGMDAVPGPVTRLMLMTTLTGVSIALFP 300
300 VAQAPTALVALAVAYGFTSGALAPLAFSVLPGLIGTRIRYCGLLQMIIESIGLLGPPL 359
301 VAQAPTALVALAVAYGFTSGALAPLAFSVLPGLIGTRIRYCGLLQMIIESIGLLGPPL 360
360 GYLRLDVSGNYTASFVAVAGAFLLSGGILLTLPHFFCFSTTSGPQDLVTEALDTKVPLP 419
361 GYLRLDVSGNYTASFVAVAGAFLLSGGILLTLPHFFCFSTTSGPQDLVTEALDTKVPLP 420
Y 420 KEGLE 424
421 KEGLE 425

ESULT 9
S-10-072-012-302
Sequence 302, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernen, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Paturajan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigar, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517

PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remainng Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 302
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-302

Query Match 91.9%; Score 2142; DB 4; Length 426;
Best Local Similarity 99.1%; Pred. No. 9.9e-178;
Matches 420; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 MARRTEPPDGGWVVLVSFAFFQSALVFGVLRSEGVFVEFVAAAFEEQAAARVSWIASIGI 60
1 MARRTEPPDGGWVVLVSFAFFQSALVFGVLRSEGVFVEFVAAAFEEQAAARVSWIASIGI 60
61 AVQFGSPVGSALSTKFGPRPVMTGGILAAAGMLASFATSLTHLYLSIGLSGSWAL 120
61 AVQFGSPVGSALSTKFGPRPVMTGGILAAAGMLASFATSLTHLYLSIGLSGSWAL 120
121 TFAPTLACLSCTFSRRSLATGLALTGVGLSSFTFAPFQWLLSHYAMRGSLLVSALS 180
121 TFAPTLACLSCTFSRRSLATGLALTGVGLSSFTFAPFQWLLSHYAMRGSLLVSALS 180
181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYLHVAH 240
181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYLHVAH 240
241 LQDLMDPLPAAFLLSVVAISDLVGRVSWGMDAVPGPVTRLMLMTTLTGVSIALFPV 300
241 LQDLMDPLPAAFLLSVVAISDLVGRVSWGMDAVPGPVTRLMLMTTLTGVSIALFPV 300
301 AQAPTALVALAVAYGFTSGALAPLAFSVLPGLIGTRIRYCGLLQMIIESIGLLGPPLS 360
301 AQAPTALVALAVAYGFTSGALAPLAFSVLPGLIGTRIRYCGLLQMIIESIGLLGPPLS 360
361 GYLRLDVSGNYTASFVAVAGAFLLSGGILLTLPHFFCFSTTSGPQDLVTEALDTKVPLPK 420
361 GYLRLDVSGNYTASFVAVAGAFLLSGGILLTLPHFFCFSTTSGPQDLVTEALDTKVPLPK 420
Y 421 EGLE 424
421 EGLE 424
Db 421 EGLE 424

RESULT 10
US-10-072-012-304
Sequence 304, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernen, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Paturajan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha

APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 304
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-304

Query Match 91.9%; Score 2142; DB 4; Length 426;
Best Local Similarity 99.1%; Pred. No. 9.9e-178;
Matches 420; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOARVSWIASIGI 60
1 MARRTEPPDGGWGVVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOARVSWIASIGI 60
61 AVOQFGSPVGSALSTKFGPRPVMTGGLAALGMLLASFATSLTHLYLSIGLSSGAWAL 120
61 AVOQFGSPVGSALSTKFGPRPVMTGGLAALGMLLASFATSLTHLYLSIGLSSGAWAL 120
121 TFAPTLACLSCYFSRRSLATGALTGVLSSFTFAPFQWLISHYAWRGSLLVSAISL 180
121 TFAPTLACLSCYFSRRSLATGALTGVLSSFTFAPFQWLISHYAWRGSLLVSAISL 180
181 HUVACGALLRPSPSLAEDPAVGSPRAQLTSLHGPFLRYTVALTLINTGYFIPLYLVAH 240
181 HUVACGALLRPSPSLAEDPAVGSPRAQLTSLHGPFLRYTVALTLINTGYFIPLYLVAH 240
241 LODLDWDPLPAFLSLVAISDLVGRVSWGIGDAVPGPVTRLMLMTTLTGVSALFPV 300
241 LODLDWDPLPAFLSLVAISDLVGRVSWGIGDAVPGPVTRLMLMTTLTGVSALFPV 300

QY 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGILLQWIESIGLGPPLS 360
Db 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGILLQWIESIGLGPPLS 360
QY 361 GYLRDVSGNYTASFVVGAFLLSGSGLLTLPHFFCFSTTSGPODLVTEALDTKVP LPK 420
Db 361 GYLRDVSGNYTASFVVGAFLLSGSGLLTLPHFFCFSTTSGPODLVTEALDTKVP LPK 420
QY 421 EGLE 424
Db 421 EGLE 424

RESULT 11
US-10-072-012-24
Sequence 24, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Shinkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangoli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosse, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 447
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-24

```
Query Match      39.9%; Score 930.5; DB 4; Length 447;
Best Local Similarity 45.9%; Pred. No. 3.2e-72;
Matches 206; Conservative 67; Mismatches 145; Indels 31; Gaps 8
```

y 7 PPDGCGWVVVLSAFFQSALVFGLRSPGFVEFVAFAFEEOARVSMTASIGIAVQDFG 66
||||| : | : | : | : | : | : | :
c 8 PPDCGWGVVAAAFAINGLSYGLLRSLGLAFPDLAEHFDRAQDTWTISALAIVQDA 67

y 67 SPVGSAIATKEGPRPVMTGGILPAALGMLLASFATSPLHYLSIGLSSGWALLTPAFTL 126
||| ||| : | : | : | : | : | : | :
b 68 SPVGSALSTRWGARPVMVWGVLASLGWFVSAFASDILLHLVLGLLAGFGVALVPAPAL 127

y 127 ACLSCYFSRRRLATGTALTGVGLSFTTFAFFOWMLSHYAMRGSLLVASLSHLVACG 186
||| ||| | | | | | | | | | : | : | : | : | : | : | :
b 128 GTLSRYFSRRRVLAAGLALTGNASSLLPALQLLDITFGWRGALLLGAILTLPTCG 187

y 187 ALLRPSSLAEADPAVGGPRAQLT---SLTHGPFLRYTVALLINTGYFIPLYHLVAHLQ 242
||| | | | | : | : | : | : | : | : | :
b 188 ALLPLVLPGDP -APRPSPLAALGQSLFTTRAFSTIALGTALVGGYFPVYLAPHAL 246

y 243 DLDDMDLPAPAFLLSVVAISDLVGRVSWGWDGAVPCBPVTRLMLMNTLTGVS L--ALFP 299
| : | : | : | : | : | : | : | : | : | : | : | : | :
b 247 DRGLGGYGALVVAAMGDAGARLVCGWLADQGWPVLRLLAVFALTGLEWMVGLVP 306

y 300 VAOAPT A----LVALAVAYGETSGALAPLAFSVLPPELICTRRIYCGLLOWMESIGLL 355
| : | : | : | : | : | : | : | : | : | : | : | : | :
b 307 VVGGEESWGPFLLAAAVAYGLSAGSYAPLVFVGVLPGLVGVGVQATGLVMMMLSLGGL 366

y 356 GPPLSGYLRDVSGNYTASFVVAAGFLSSGILLTPHFCCSTTTSGPODVTETALDK 415
||| ||| : | : | : | : | : | : | :
b 367 GPLPLSGPLRDETGDTFASFLSSGLILSSGFITYIGLPR-----ALPSCGP-----ASPBA 416

y 416 VP LPKEGLEGGLNSTSGPESQS-LTAPG 443
417 TPPEETG-----ELPPAQAVLLSPG 437

.RESULT 12
.S-10-072-012-30
Sequence 30, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchernev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patlurajan, Meera
APPLICANT: Shimkets, Richard Li, Li
APPLICANT: Gangoli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsobrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514

```

; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ.ID NO 30
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-30

Query Match      39.5%; Score 920; DB 4; Length 472;
Best Local Similarity 45.8%; Pred. No. 2.8e-71;
Matches 206; Conservative 67; Mismatches 145; Indels 32; Gaps 9;

QY      7  PPDGWGRVVVLSAFFQSALVEGVLSFGVFFVEFVAAFEEOAARVSWIASIGIAVQOFG 66
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      32  PPDGCGWVVAATAAFAINGLSYGLRSGLAFPDLAEHFDRSAQDTAWISALALAVQQA 91

QY      67  SPVGSALSTKFGPRPVMTGILAAQMLLASFATSLTHLYLSIGLISGSGWALTFAPTL 126
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      92  SPVGSALSTKFGARPVVMVGVLASIGFVESAFAASDLLHLYLGLLAGFGWALVFAPAL 151

QY      127  ACLSCYFSRRSLATGLALTGVLSFTFAFFQWLLSHYAWRGSLLVLSLHLVACG 186
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      152  GTLSRYFSRRVLAVGLALTNGASLLAPALQLLDFTFGWRGALLLGAITLHLP CG 211

QY      187  ALLRPPSLAEDPAVGGPRAQLT---SLHNGPFLRYTVALTLINTGYEIPYHLVAHLQ 242
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      212  ALLPLVLPGDDP-APPRSPALALGLSLFTRRAFSIFALGTALVGGYFVPYVHLAPHAL 270

QY      243  DLDWDPLPAAPFLISVAISDLVGRVYSGWLGDVAPGPVTRLLMWTTLTGVS L--ALFP 299
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      271  DRGLGGYGAALVVAAMGDAGARLVCGMLADQGWVPLPRLAVFGALTGLGWVGLVP 330

QY      300  VAQAFTA---LVALAVAYGFTSGALAPLAFSVLPPELIGTRIRYCGGLQIMIESIGLL 355
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      331  VVGGEESWGGPPLAAAVAYGLSAGSYAPLVFGVLPGLVGVGVQATGLVMMLMSLGLL 390

QY      356  GPPLSG-YLRDV$GNNTASFSVAVAGAFILSGSGLTLTPHFFCFSTTSGPQDLVTEALDT 414
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      391  GPPLSGKFLRDETGDFTASFLSGSLISGSFYIYGLPR---ALPSCGP-----ASPP 440

QY      415  KVPLPKEGLEGGLNSTESGPESQS-LTAPG 443
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      441  ATPPETG-----ELLPAQAVALLSFG 462

RESULT 13
US-10-644-807-224
; Sequence 224, Application US/10644807
; Publication No. US20060057582A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 83 Human Secreted Proteins
; FILE REFERENCE: PS735
; CURRENT APPLICATION NUMBER: US/10/644, 807
; CURRENT FILINGDATE: 2003-08-21

```

PRIOR APPLICATION NUMBER: PCT/US02/05064
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,658
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/304,444
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 445
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 224
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-644-807-224

Query Match 30.2%; Score 704; DB 5; Length 142;
Best Local Similarity 99.3%; Pred. No. 4.1e-53;
Matches 139; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 285 MLWTTLTGVSIALFPVQAAPTALVALAVAYGFTSGALAPLAFSVLPPELIGTRRIYCGLGL 344
Db 1 MLWTTLTGVSIALFPVQAAPTALVALAVAYGFTSGALAPLAFSVLPPELIGTRRIYCGLGL 60
QY 345 LQMIESTIGLGPPLSGYL RDVSGNYTASFVAVAGAFLLSGSGLLTLPHFPCFSTTTSGP 404
Db 61 LQMIESTIGLGPPLSGYL RDVSGNYTASFVAVAGAFLLSGSGLLTLPHFPCFSTTTSGP 120
QY 405 QDLVTEALDTKVPLPKEGLE 424
Db 121 QDLVTEALDTKVPLPKEGLE 140

RESULT 14
US-10-644-807-314

; Sequence 314, Application US/10644807
; Publication No. US20060057582A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 83 Human Secreted Proteins
; FILE REFERENCE: PS735
; CURRENT APPLICATION NUMBER: US/10/644,807
; PRIOR FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: PCT/US02/05064
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,658
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/304,444
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 445
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 314
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-644-807-314

Query Match 30.2%; Score 704; DB 5; Length 142;
Best Local Similarity 99.3%; Pred. No. 4.1e-53;
Matches 139; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 285 MLWTTLTGVSIALFPVQAAPTALVALAVAYGFTSGALAPLAFSVLPPELIGTRRIYCGLGL 344
b 1 MLWTTLTGVSIALFPVQAAPTALVALAVAYGFTSGALAPLAFSVLPPELIGTRRIYCGLGL 60
Y 345 LQMIESTIGLGPPLSGYL RDVSGNYTASFVAVAGAFLLSGSGLLTLPHFPCFSTTTSGP 404
b 61 LQMIESTIGLGPPLSGYL RDVSGNYTASFVAVAGAFLLSGSGLLTLPHFPCFSTTTSGP 120
Y 405 QDLVTEALDTKVPLPKEGLE 424
b 121 QDLVTEALDTKVPLPKEGLE 140

ESULT 15

US-09-829-432-4
; Sequence 4, Application US/09829432
; Publication No. US2003016522A1
; GENERAL INFORMATION:
; APPLICANT: KETCHUM, Karen et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001013
; CURRENT APPLICATION NUMBER: US/09/829,432
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 09/815,301
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/254,554
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-829-432-4

Query Match 27.7%; Score 645; DB 3; Length 456;
Best Local Similarity 36.1%; Pred. No. 2.4e-47;
Matches 170; Conservative 70; Mismatches 193; Indels 38; Gaps 9;

QY 3 RTEPPDGGWGRVVLSAFQSALVFGLRSFGVFVEFVAAFEEQARVSWIASIGIAV 62
Db 1 RGAGPPDGGWVVLGACFVITGAFYGFPAKAVSVFRELKRDGAGYSDTAWVSSIMLAM 60
QY 63 QQFSPVGSALSTKFGPRPVMTGILALGMLASFATSLTHLYSLGSSGWAITF 122
Db 61 LYGTGPLSSILVTRFGCRPVMLAGGLASAGMILASFSARLLELYTAGVLTGLALNF 120
QY 123 APTLACLSYFSRRSLATGLALTGVGLSSFTFPAFFQWLLSHYAWRGSLIVSALSLH 182
Db 121 QPSLIMGLYFERRRPLANGLAAGSPVFLSTLSPGLQLLGERFGWRGFLFGGLLHC 180
QY 183 VACGALLRPPSLAE---DPAVGPPRAQLTSLH-----HGFRLRYTVALTLINTGYFIY 234
Db 181 CACGAVMRPPPGQPRDPAPPGGARHRQLDLAVCTDRTFMVYMTKFLMALGLFVPA 240
QY 235 LHLVAHLQDLMDPLPAFLLSVVAISDLYGRVSG---WLGDAVPGVTRLLMLWTTLT 291
Db 241 ILLVNIAKADAGVDAEAAFLSIVGFVDIVARPACGALAGLRLRP-HVPYLFSLALLAN 299
QY 292 GVSIALFPVQAAPTALVALAVAYGFTSGALAPLAFSVLPPELIGTRRIYCGLQMIESI 351
Db 300 GLTDLISARARSYGTLVAFCIAGLSYGMVGAHQFEVLMATVGAPRPPSALGLVLVEAV 359
QY 352 GGLGPPPLSGYL RDVSGNYTASFVAVAGAFLLSGSGLLTLPHFPCFSTT----- 400
Db 360 AVLIGPPSAGRLVDALKNYEILFYLAGS-EVALAGVEMAVTTYCCLRCSKNISSGRSAEG 418
QY 401 -TSGPQDLVTEALDTKVPLPKEGLEGLNSTESGPESQSLTAPGLLPRLG 450
Db 419 GASDPEDV--EAERDSEPMFA-----STE--EPGSLEALEVLSPRAG 456

Search completed: September 7, 2006, 13:12:06
Job time : 184 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2006, 13:09:11 ; Search time 34 Seconds

(without alignments)
944.240 Million cell updates/sec

Title: US-10-654-428-2

Perfect score: 2331

Sequence: 1 MARRTEPPDGGWGRVVLSA.....SLTAPGLLPRLGHRTPV 457

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 254368 seqs, 70249769 residues

Total number of hits satisfying chosen parameters: 254368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA New:*

1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB.pep:*

2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB.pep:*

3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB.pep:*

4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*

6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB.pep:*

7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB.pep:*

8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	591.5	25.4	465	6	US-10-553-520-168 Sequence 168, App
2	386.5	16.6	477	6	US-10-530-837-8 Sequence 8, Appli
3	236	10.1	673	7	US-11-251-208-168 Sequence 168, App
4	224.5	9.6	457	6	US-10-449-902-38760 Sequence 38760, A
5	157.5	6.8	628	6	US-10-449-902-42609 Sequence 42609, A
6	150.5	6.5	624	6	US-10-449-902-33879 Sequence 33879, A
7	150	6.4	591	6	US-10-449-902-48618 Sequence 48618, A
8	147	6.3	529	7	US-11-174-307B-3218 Sequence 3218, Ap
9	140.5	6.0	549	7	US-11-274-317A-2 Sequence 2, Appli
10	140	6.0	508	7	US-11-296-657-3 Sequence 3, Appli
11	138	5.9	550	7	US-11-056-355B-83755 Sequence 83755, A
12	138	5.9	558	7	US-11-056-355B-83754 Sequence 83754, A
13	138	5.9	577	7	US-11-056-355B-83753 Sequence 83753, A
14	137.5	5.9	470	7	US-11-330-403-1885 Sequence 1885, Ap
15	136.5	5.9	457	7	US-11-330-403-8385 Sequence 8385, Ap
16	136	5.8	591	6	US-10-449-902-43177 Sequence 43177, A
17	135.5	5.8	603	6	US-10-449-902-33462 Sequence 33462, A
18	135.5	5.8	603	7	US-11-056-355B-19289 Sequence 19289, A
19	134.5	5.8	519	6	US-10-449-902-35276 Sequence 35276, A
20	133.5	5.7	329	6	US-10-449-902-35276 Sequence 46252, A
21	133	5.7	365	7	US-11-299-311-1 Sequence 1, Appli
22	133	5.7	547	6	US-10-474-894-1 Sequence 1, Appli
23	133	5.7	558	6	US-10-953-349-33463 Sequence 33463, A
24	133	5.7	558	7	US-11-056-355B-19290 Sequence 19290, A
25	132.5	5.7	471	7	US-11-056-355B-107639 Sequence 107639, A

26	132.5	5.7	471	7	US-11-056-355B-118878	Sequence 118878, A
27	132.5	5.7	483	7	US-11-174-307B-5500	Sequence 5500, Ap
28	132.5	5.7	483	7	US-11-056-355B-107638	Sequence 107638, A
29	132.5	5.7	483	7	US-11-056-355B-118877	Sequence 118877, A
30	132.5	5.7	541	7	US-11-056-355B-107637	Sequence 107637, A
31	132.5	5.7	541	7	US-11-056-355B-118876	Sequence 118876, A
32	132	5.7	514	6	US-10-449-902-40710	Sequence 40710, A
33	131.5	5.6	499	6	US-10-953-349-33464	Sequence 33464, A
34	131.5	5.6	499	7	US-11-056-355B-19291	Sequence 19291, A
35	130	5.6	443	7	US-11-056-355B-85631	Sequence 85631, A
36	130	5.6	444	7	US-11-056-355B-85630	Sequence 85630, A
37	130	5.6	537	7	US-11-056-355B-85629	Sequence 85629, A
38	128	5.5	510	6	US-10-449-902-36204	Sequence 36204, A
39	127.5	5.5	516	7	US-11-330-403-763	Sequence 763, App
40	126	5.4	388	6	US-10-471-571A-3466	Sequence 3466, Ap
41	126	5.4	461	7	US-11-330-403-2630	Sequence 2630, Ap
42	125.5	5.4	534	6	US-10-953-349-22703	Sequence 22703, A
43	124	5.3	461	7	US-11-330-403-10052	Sequence 10052, A
44	123	5.3	397	6	US-10-471-571A-5132	Sequence 5132, Ap
45	122.5	5.3	439	7	US-11-330-403-15593	Sequence 15593, A

ALIGNMENTS

RESULT 1

US-10-553-520-168

Sequence 168, Application US/10553520

Publication No. US20060188885A1

GENERAL INFORMATION:

APPLICANT: Bodian, Dale

APPLICANT: Daouti, Sherif

APPLICANT: Kumar, Chandrika

APPLICANT: Latario, Brian

APPLICANT: Quintavalla, Joseph

TITLE OF INVENTION: High throughput functional genomic

TITLE OF INVENTION: Screening methods for osteoarthritis

FILE REFERENCE: 4-33178

CURRENT APPLICATION NUMBER: US/10/553.520

CURRENT FILING DATE: 2005-10-14

PRIOR APPLICATION NUMBER: 60/463,933

PRIOR FILING DATE: 2003-04-18

NUMBER OF SEQ ID NOS: 180

SOFTWARE: FastSeq for windows Version 4.0

SEQ ID NO 168

LENGTH: 465

TYPE: PRT

ORGANISM: homo sapiens

US-10-553-520-168

Query Match 25.4%; Score 591.5; DB 6; Length 465;

Best Local Similarity 33.7%; Pred No. 1e-37;

Matches 141; Conservative 76; Mismatches 193; Indels 9; Gaps 4;

QY PDGWMGRVVVLSAFFQSALVFGVLSFGVFFVEFVAFAFEQARVSWIASIGIAVQOFGS 67

DB PDGWMGMAVLFGCFTVITGFSYAFAFKAVSVFFKELIGFEGIGYSDTAWISSILLAMLYGTG 75

QY PVGSALSTKFGPRPVVMTGILALGMLASFATSLTHLYLSIGLSGSGWALTFAPTLA 127

DB PLCSVCVNRFGCRPMVHVGFLFASLGMVAASFCSIITQVYLLTGTITGLALNFQPSLI 135

QY CLSCYFSRRSLATGIALTGVLSSFTFAPFQWLSHYAMRGSLLVSAISLHVACGA 187

DB MUNKRYFSKRRPMANGIAAAGSPVFLCALSPGOLLQDRYGMRGFLLIGLLNCCVCAA 195

QY LURPSLAEDPAVGGPRAQL---TSLHNGPFLRYTVALTLLINTGYFTPLYLHV AHLQD 243

DB LMRPLVVTAAQPGSGPPRPSRRLLDLSVFRDRGFVLVAVAASVMVLGLFVPFVVSYAKD 255

QY LDWDPLPAFLLSVAISDLVGRVSGW---LGDVAVPGVTRLLMLWTLLTGVSIALFPV 300

DB LGVPDTKAFLTLITLIGFIDIFARPAAGFVAGLGKVRPYSV-YLFSPFMFNGLADLAGST 314

1 AQAPTALVALAVAYGFTSGALAPLAFSVLPPELIGTRRIYCGILQMIIESIGLLGPPLS 360
2 ACIDYGLVVFCEIFFGISYGMVGAQLQFEVLMAIVGTHKFSSAIGLVLLMEAVAVLVGPESG 374
3 GYLKRVSGNYTASFVAVAGAFLLSGGILLTLPHFECFSTTSGPQDLVTEALDTKVPPLP 419
4 GKLLDATHVVMYVFILAGAEVLTSILL-LGNFECIRKKPKPEQPEVAAAEKEKLHKP 432

350LT 2
3-10-530-837-8
Sequence 8, Application US/10530837
Publication No. US20060168670A1
GENERAL INFORMATION:

APPLICANT: LG Life Sciences Ltd.
TITLE OF INVENTION: Gene Familillies Associated With Cancers
FILE REFERENCE: L103PCT006
CURRENT APPLICATION NUMBER: US/10/530,837
PRIOR APPLICATION NUMBER: 2005-04-08
PRIOR FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: US60/419911
PRIOR FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: US60/419912
PRIOR FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: US60/420088
PRIOR FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: US60/434243
PRIOR FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: US60/434278
PRIOR FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: US60/438278
PRIOR FILING DATE: 2003-01-03
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Kopatentcin 1.71
SEQ ID NO 8
LENGTH: 477
TYPE: PRT
ORGANISM: Homo sapiens
S-10-530-837-8

Query Match 16.6%; Score 386.5; DB 6; Length 477;
Best Local Similarity 24.1%; Pred. No. 5e-22;
Matches 116; Conservative 76; Mismatches 150; Indels 139; Gaps 10;

Y 8 PDGCGRRVVVLSAFQOSALVFGVLSFGVFVEFVAAFEEOQARVSWIASIGIAVQOFGS 67
b 7 PDGCGWGVIVFVSFLTQFLCYGSPLAVGVLYTEWLDAFGEKGKTAWVGSILASGVGLLAS 66
Y 68 PVGSALSTKFGPRPVMTGILALGMLASFATSLTHLYLSIGLISGSWALTFAPTLA 127
b 67 PVCSLCVSSFGARBPVTIFSGFMVAGLMSFAPNIFYLFFSYGIIVGS----- 115
Y 128 CLSCYFSRRRSLATGLALTGVGLSFTFAPFQWLLSHYAMRGSLLVSAISLHLVACGA 187
b 116 -----SVGL--FITYALQRMVLEFYGLDGCCLLVGALALNITLACGS 154
Y 188 LLRPPSLAEDPA-----VGGPRAQLT----- 208
b 155 LMRPLQSSDCPLPKKIAPEDLPDKYSIYNEKGNLEENINILDKSYSEKCRITLANGD 214
Y 209 -----SLUHGPFLLRYT----- 220
b 215 WKQDSLHLKHPVTYHTKEPETYKKKVAEQGYFCKQLAKRKWQLYKNYCGETVALFKNVF 274
Y 221 ----VALTLINTGYFIPLYHL--VAHLQDLMDP--LPAFFLSVVAISDLVGRVSVGWL 272
b 275 SALFIALILFDIGFPPSLMEDVARSSNVKEEFIMP--LISTIGMTAVGKLLLGIL 331
Y 273 GDAVPGPVTRLMLMTTLTGVSIALFPVAQAPTALVALAVAYGFTSGALAPLAFSVLP-- 330
b 332 ADFKWINITLVYVATLLIMGLALCAIPFAKSVVTLALLSGILGFLTG-----NMSIFPYV 386
Y 331 --ELICTRRRIYCGI GLIQMTESIGLLGPPLSGYLRDVSGNVTASFVAVAGAFLLSGSGL 388

Db 387 TTKTVGIEKLAHAYGILMFAGLGNISGPPIVGWFYDWTQTYDIAFYFSGFCVLLGGFIL 446
QY 389 L 389
Db 447 L 447

RESULT 3
US-11-251-208-168
Sequence 168, Application US/11251208
Publication No. US20060137043A1
GENERAL INFORMATION:

APPLICANT: Puzio, Piotr
APPLICANT: Chardonens, Agnes
APPLICANT: Shirley, Amber
APPLICANT: Wang, Xi-Qing
APPLICANT: Sarrila-Millan, Rodrigo
APPLICANT: Mckersie, Bryan
APPLICANT: Chen, Ruoying
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING PROTEINS ASSOCIATED WITH ABIOTIC
TITLE OF INVENTION: STRESS RESPONSE AND PLANT CELLS AND PLANTS WITH INCREASED
TITLE OF INVENTION: TOLERANCE TO ENVIRONMENTAL STRESS
FILE REFERENCE: 13311-00015-US
CURRENT APPLICATION NUMBER: US/11/251,208
PRIOR FILING DATE: 2005-10-14
PRIOR APPLICATION NUMBER: PCT/US2004/011888
PRIOR FILING DATE: 2004-04-15
PRIOR APPLICATION NUMBER: EP 03008080.8
PRIOR FILING DATE: 2003-04-15
PRIOR APPLICATION NUMBER: EP 03009728.1
PRIOR FILING DATE: 2003-05-02
PRIOR APPLICATION NUMBER: EP 03016672.2
PRIOR FILING DATE: 2003-08-01
PRIOR APPLICATION NUMBER: EP 03022225.1
PRIOR FILING DATE: 2003-09-30
NUMBER OF SEQ ID NOS: 852
SOFTWARE: Patentcin version 3.3
SEQ ID NO 168
LENGTH: 673
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-11-251-208-168

Query Match 10.1%; Score 236; DB 7; Length 673;
Best Local Similarity 24.5%; Pred. No. 2.4e-10;
Matches 107; Conservative 69; Mismatches 213; Indels 48; Gaps 14;

QY 7 PPDGCGRRVVVLSAFQOSALVFGVLSFGVFVEFVAAFEEOQAR--VSWIASIGIAVQ 63
Db 206 PPDGCGWVVTFCVFLTMFSTWGCNASFVDLAYVYNHDTYRPGASKYDVALIAGLTVFLG 265
QY 64 QFGSPVGSALSTKFGPRPVMTGILALGMLASFATSLTHLYLSIGLISGSWALTFA 123
Db 266 QLLSPLVMALMRIIGLRTMLFGDAVMLAAVYLLASFSTTKLMQLYVTQGFVWGCSISLTFV 325
QY 124 PTLACLSYFSRRRSLATGLALTGVGLSFTFAPFQWLLSHYA--WRGSLLVSAISL 180
Db 326 PATTVLPGWFLKKRAVAMGVSLLTGAGGVVYGLATNKMLSDFGNTRWCLRIIGISC-SI 384
QY 181 HUVACGALL--RPPSLAEDPAVG--GPRAOULTSL-----LHHGPF-----LRYTVALT 224
Db 385 SVLVAIALALKERNPT----PAIGLKSPRAMFEQKAMFSLKVTITKPFVLLIALWMFALF 440
QY 225 LINTGYFIPLYHLVAH-LQDLMDPPLPAAFFLSVVAISDLVGRVSVGWLGDVPGPVTRL 283
Db 441 AYNNMVFTLLSSYAISKGSJSHD-----ASTLTAIILNGSQSIGRPLMGLAGDKF--GRANVT 494
QY 284 LMLMTTLTGVSIALFPVAQAPTALVALAVAYGFTSGALAPLAFSVLPPELIGTRRI----- 338
Db 495 IVLTTLLTYMFAFWIPAHTFVQLIFFSILVSGCVGVANVMNTVLIADMWKBEFPLPAWA 554
QY 339 ---YCGLGLLQMIIESIGLLGPPLSGYLRDVSGNVTASFVAVAGAFLLSGSGLILLTLPHF 395

Db 555 FVNYCGAPFLVCEVIAQALIVE-----KDKSNPYLHAQIFCGCCFIAL-ILISIREY 608
QY 396 CFSITTSQPDLVTEAL 412
Db 609 SIRMKLTERQAMTNEKL 625

RESULT 4

US-10-449-902-38760
; Sequence 38760, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38760
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-38760

Query Match
Best Local Similarity 9.6%; Score 224.5; DB 6; Length 457;
Matches 110; Conservativity 25.5%; Pred. No. 1.2e-09;
Matches 110; Mismatches 62; Indels 169; Gaps 19;

QY 7 PPDG-GW---GRVVVLSAFQOSALVFGVLSFGVFVEFVAAFEQAAVSWIASIGIAV 62
Db 43 PPEGYGWVCGAVSLINGF-----TWGVAASYGYLSYSSSHFPTATPMDYAFIGLI 97
QY 63 QQFG-----SPVGSALSTKFGPRPVMTGG-ILALGMLLASFATSLTHLYLSIGLGS 116
Db 98 --FGCALFSPVWTLTRELG-RKLVMTGCGFLMAAGFVAASFATQPMQLYLSQGAAVGT 154
QY 117 GWALTFAPFLACLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLSHYAWRGSLLVS 176
Db 155 GMAIFTPSVQVWPQWFLRRSLAGTASAGSGFGGIAFA-----LATDAMIRQLSLAW 208
QY 177 ALSHLVAC-----GALLRPPSLAEDPAVGGPRAQLTSLHGHFLRYTVA 222
Db 209 SFRITGICCLIGNLIGTALIRDRNAQIKPPQL-----GFATHLLK-----RYDCL 253
QY 223 LTL-----INTGYFIPYLHLVAH--LQDLWDPLPAAFLLSVVAISDLVGRVSGWUGD-- 274
Db 254 LLLAWAFTNIGYIVILYSVSSYAVQVAGLTOKQAGILTAMLMGTGIGRAPAGLVSDRF 313
QY 275 ---AVPGPVT-----RLMLWTTLTGVSALFPVAQAPTALVALAVAGFTSGALAPLAF 326
Db 314 GRIQVAAVITLACAIISVFAIWIPTSYGVLIF-----FSLVSGAILGVYW 358
QY 327 S-----VLPELIGTRRIYCGLLQMIIESIGLLGPPLSGYLR--DVS--GNYTASFVVA 379
Db 359 TCVGPPLCAETIAGLKEVPSFLSLHMLTVLPTTFAEVIALYLRRPDMGRWGYLYTQLFTGF 418
QY 380 FLTSGSGILLTL 391
Db 419 AYLVAAGFLFEL 430

ESULT 5
S-10-449-902-42609
Sequence 42609, Application US/10449902

; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42609
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-42609

Query Match
Best Local Similarity 6.8%; Score 157.5; DB 6; Length 628;
Matches 97; Conservativity 21.7%; Pred. No. 0.00022;
Matches 97; Mismatches 65; Indels 171; Gaps 19;

QY 26 LVFGVLSRF---GVFEVE-FVAAFEQAAAR---VSWIASIGIAVQFGSPVGSALSTK 76
Db 165 IVLGLKGFVGSGLAIFTQLYVAIYGDDAKSLVLLIAMP-----AAISIL 210
QY 77 FGPRPVMTGILAL-----GMLASFATS---LTHLYLSIGLLSGSWALTFAP 124
Db 211 F-----VHTVRIMPYLPSRRRRADGELIASATSNDAFFCFELYISIAL-----ATYLL 258
QY 125 TLACLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLSHYAWRGSLL----- 172
Db 259 TMIWVQ---NQTNFSHTAYVVSATALLLVLFLLVNVVVIKQEVQIKKELDDSLREPTVTI 315
QY 173 --LVVLSLSHLVACGALLRPPSLAEDPAVGGPRAQLTSLH--GPFLR--YTVALTL 225
Db 316 EKPAALAAQMOSAITTKPKTETPS-SSSPAPAPSPCCSGCKMHFNPAPQGEDYTTIQAL 374
QY 226 INTGYFIPYLHLVAHL-----QDLWDPLPAAFLLSVVAISDLVGRVSGW 271
Db 375 VSVDMVLFLATIGVGGTLLTAIDNMGQIGSLGYPAKSIKFTISLSIMNYAGRVTSGF 434
QY 272 LGD-----AVPGPVTRLMLWTTLTGVSALFPVAQAPTALVALAVAGFTSGALAPLA 325
Db 435 ASEMFLARYRFPRLMLTAVLLLACVGHLLIAFGVAQ---SLYASVIGFCFGAQWPLL 491
QY 326 FSVLPFLIGTRRIYCGLLQMIIESIG-LLGPPLSGYLRDV-----SGNYTA- 372
Db 492 FAIISSEVFGKXYSTLYNPGSVASPVGAYVLNVRVAGYLYDVEAARQHGSLAGGDKTCL 551
QY 373 -----SFVVAAGAFLLSGSGLTL 391
Db 552 GVQCFFRAFLIITATVAGALISLVL 577

RESULT 6
US-10-449-902-33879
; Sequence 33879, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269

PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: JP 2002-383870
PRIOR FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 56791
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 33879
LENGTH: 624
TYPE: PRT
ORGANISM: Oryza sativa
US-10-449-902-33879

Query Match 6.5%; Score 150.5; DB 6; Length 624;
Best Local Similarity 19.0%; Pred. No. 0.00075;
Matches 110; Conservative 65; Mismatches 151; Indels 253; Gaps 22;

```

Y 2 ARTEPPDGGWG---RVVLSAFFQSALVFGVLRSGVFVEFVAAFEQAAARVSWI--- 55
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 16 AAPSPPP--WQNPLRVILTAFAKQAV-----GRWFTVFASLLILTAGGATYIFGI 66
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 56 -----ASIG-----IAVQFGSPVGSALSTKFG-----PRPVMTGGILALGMLLA 97
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 67 YSPALKASLGYPDHTLNTVSEFKDGLANGLVSLGLINEVTPPVVLAIGAMNLSGYLMV 126
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 127 YLAAGRTAAPVWLVCIFYEVGANSQS-----FANTGALVTCVKNFPESRGVILGILKG 181
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 147 GVGLSSFTFA-----147-----156
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 182 FVGLSGAVYTYQLYLAFYGDPAKSLILLIAWLPAAVSVVEVHTVRIMPRRRRGQETSVD 241
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 157 PFP-----QWLLSHYAW---RGSLLLVSALSLLHV-----183
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 242 PFFCELVISIGLAYLLVMIVQKQAFSRTAYSCAAAMLLIVLPLCVVIKQEFKIR 301
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 184 -----ACGA-----184-----204
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 302 ERLELAALAPPHITITVLEMSKETERSPRSSPAPAEYSWYKGMFRPPARGED----- 354
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 205 AQLTSLHHGPFRLRYTVALTLINTGYFIPLYHLVAHL-----QDLWDPLP 250
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 355 -----YTILQALVSVMNAVLFVATICGVGGLTAIDNMGOIGQSLGYPARS 400
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 251 AAFLLSVVAISDLVGRVSGMLGDA-----VGPVTRLMLMTTLTGVSLLA-----LFP 299
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 401 TMTFVSLISIMWYAGRVAAAGFASAEFVERWRLLRPL-----VLTGILLACAGHLLI 452
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 300 VAOAPTALVALAVAGFTSGALAPLAFSVLPPELIGTRRIYCGILQIMESIGLGPPL 359
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 453 ALGVPRALYAASVVIQFCFGAQWPLVFAIISEVF-----GLKYSTLYNFGMASP-- 503
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 360 SGYLRDVSGNYTASFVVA-----GAFLLSGSG 386
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 504 -----VGSYILNVLVAGRLYDAEAGRQPGAGLAAGAG 535
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 7
US-10-449-902-48618
; Sequence 48618, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11

NUMBER OF SEQ ID NOS: 56791
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 48618
LENGTH: 591
TYPE: PRT
ORGANISM: Oryza sativa
US-10-449-902-48618

Query Match 6.4%; Score 150; DB 6; Length 591;
Best Local Similarity 22.9%; Pred. No. 0.00077;
Matches 99; Conservative 64; Mismatches 161; Indels 108; Gaps 20;

```

Y 26 LVFGVLRSGVFVEFVAAFEQAAARVS-WIASIGIAV-----QQFGSPVGSALST---KF 77
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 85 LAIGAMNULGGYLMYLSVTGRVGARTPLMLVCLYIYVAGANSQAFANT--GALVTCVKNF 142
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 78 GPRPVMTGGILALGMLLASFATSLTHLYLSIGLSGS-----GW-----ALTEFA 123
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 143 PESRGVMGLLKGFGVLSGAIF-----TQLYLSFYGGGGSNTKPLILLVGWLPAAISVAFL 198
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 124 PTLACL-----SCYFSRR-----SLATGLALTGVGLSSFTTFAPFFQWLLSHYAMRGSLL 173
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 199 GTIRIRAPRSPTAARREYRAPFCGFLVYSLALAYLLVIVLQKRFKFTRAEYAVSAAV 258
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 174 LVSALSLLHV-----ACGALLRP--PSLAEDPAVGG-----202
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 259 FAALLAPPAIVLREAAALFRKTPKREADDVPALSAATKPSAPAAETPPATAMERVVRL 318
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 203 ---PRAQ-----LTSLLHHGPFRLRYTVALTLINTGYFIPLYHLVAHL-QDLWDPLPAAFL 254
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 319 RPPRGEDYTLQALVSDVMVLLFTA--TVFGVGTLTAIDNMGOIGESLGYPQRSIATL 376
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 255 LSVVAISDLVGRVSGMLGDAV-----GPPVTRLMLMTTLTGVSLLAFPVQAQPTA 306
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 377 VSLISIMNYLGRVSAGFASDALLSRYGISRPPVVTGVLL--TVAGHLLVAFGV---PGS 431
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 307 LVALLAVAYGFTSGALAPLAFSVLPPELIGTRRIYCGLLQIMESIGLGPPLSGYLRDV 366
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 432 LYAASVLIGFCGAAYPMILAIISEVF-----GLKYSTLYNVGNVACP-----475
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 367 SGNYTASFVVAG 378
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Y 476 VGSYILNVRVAG 487
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 8
US-11-174-307B-3218
; Sequence 3218, Application US/11174307B
; Publication No. US20060143729A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY
; TITLE OF INVENTION: USEFUL FOR MODIFYING PLANT CHARACTERISTICS
; FILE REFERENCE: 2750-1601PUS2
; CURRENT APPLICATION NUMBER: US/11/174,307B
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583,671
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,781
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,651
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 5544
; SEQ ID NO 3218
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Pfam Name: MFS_1; Pfam Description: Major Facilitator Superfamily

```

; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Pfam Name: Sugar_tr; Pfam Description: Sugar (and other)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 42569720; NR Description: transporter-related
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 34908356; NR Description: P0529H11.31 [Oryza
; OTHER INFORMATION: sativa (japonica cultivar-group)] >gi|18461191|dbj|BAB84388.1|
; OTHER INFORMATION: P0529E05.3 [Oryza sativa (japonica cultivar-group)]
; OTHER INFORMATION: >gi|20805189|dbj|BAB92858.1| P0529H11.31 [Oryza sativa (japonica
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 51536596; NR Description: At3g46980 [Arabidopsis
; OTHER INFORMATION: thaliana] >gi|22331630|ref|NP_190282.2| transporter-related
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 42572593; NR Description: transporter-related
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 56784504; NR Description: putative membrane
; OTHER INFORMATION: glycoprotein HP59 [Oryza sativa (japonica cultivar-group)]
; OTHER INFORMATION: gi|56784255|dbj|BAD81937.1| putative membrane glycoprotein HP59
; US-11-174-307B-3218

```

Query Match 6.3%; Score 147; DB 7; Length 529;
Best Local Similarity 21.3%; Pred. No. 0.0012;
Matches 105; Conservative 60; Mismatches 173; Indels 154; Gaps 22;

```

; 11 GW-----GRVVLSAFQSAIVFGVLSFGVFVEFVAAFEEOARVSWIAS----- 57
; 76 GWRARRRGREVAEC---SASLEGVRHGAATAAVPSVPALPERAKVALVAAMVLLCNA 132
; 58 -----IGIAVQO-----FGSPVGSALSTKFGPRPVMTGIT 89
; 133 DRVMSVAVVPLAAGHGWSSSFVGIQSSFLMGVYVFSSWVGALADRYGCKVMAGAAL 192
; 90 AALGMLASFATSLT-HLYLSIGLSSGWMALTFAPTLACLSCYF-SRRSLATGLALTG 147
; 193 WSLATFLTPMAASQSATMLLAVRLFGVAEGVAFPTMSTFLPKWFPTHERATAVGLSMG 252
; 148 VGLS---SFTFAPFFQWLLSH-----YAMRGSLL-----VSALS 179
; 253 FHIGNVVSFLATPI---IMSHIGLAGTFAFASLGYLWLSVMLNVESDPIDSRITKSE 309
; 180 LHLVACGALLRPPSLAEDPAVGPRAOQLTSLHGHGFLRYTVALTLINTGYFI-----P 233
; 310 LQILLAGR-----SKSKVKGSKSPSLREVFSKMEMMAITVANVINNWGYFVLLSWMPV 362
; 234 YLHLV--AHLQDLDM-DLP-----AAFL---LSVVAISDLVGRVVSGL 272
; 363 YFKTVYNNVLKQAAWFSAPMGVMAISGYVAGASADFMKISGISIVRVRKIMQSI--GFI 420
; 273 GDAVPGPVTRLMLMTTLTGVSLLALFPVAQAPT-ALVALAVAAGFTSGALAPLAFSV--- 328
; 421 G-----PGVSLCLRFPAQTPSAAAVIMTAALGLSSCSQAGYFCNVQDI 463
; 329 LPELIGTRRIYCGLGLQMTESISGGLGPPLSGYLRDVSGNYTASFVVAAGAFLLSGSGL 388
; 464 APKYAGSLH-----GMTNGIGTVAIVSTVGAGYFVQMLGSFOA-----F 503
; 389 LTLPHFCFSTT 400
; 504 LTLTAVLYFSAT 515

```

;SULT 9

```

; US-11-274-317A-2
; Sequence 2, Application US/11274317A
; Publication No. US20060177912A1
; GENERAL INFORMATION:
; APPLICANT: FARMICK, Mike
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE dep34 GENE
; FILE REFERENCE: 032301WD203
; CURRENT APPLICATION NUMBER: US/11/274,317A
; PRIOR FILING DATE: 2005-11-16
; PRIOR APPLICATION NUMBER: US/09/946,763
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-11-274-317A-2

```

Query Match 6.0%; Score 140.5; DB 7; Length 549;
Best Local Similarity 21.3%; Pred. No. 0.0038;
Matches 105; Conservative 61; Mismatches 199; Indels 127; Gaps 17;

```

; QY 13 GRVVLSAFQSAIVFGVLS-----FGVFVEFVAAFEEOARVSWIASIGIAVQOFG 66
; Db 45 GRV-----GFIIALMLMLSSLGQTFISALPTIVGEL-GGVNMTWVITAFILQOTIS 99
; QY 67 SPVGSALSTKFGPRPVMTGILALALGMLASFATSLTHLYLSIGLSSGWMALTFAPTL 126
; Db 100 LPIFGKLGDFGRKYLEMFAIALFVVGSIICALAQNMNTLIVARALOGIAGGMLILSOA 159
; QY 127 ACLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLLSHYAWRGSLL-VSALSILHVA 185
; Db 160 ITADVTTARERAKYMGIMGSVFGSLISLGPLLGWFTDGPGRWGLWNVPIGIIALVAI 219
; QY 186 GALLRPP-----SLAEDPAV-----GPPRAQLTSLHGHGFLRYTVA- 222
; Db 220 AVLLKLPARERGKVSVDWLGSIFMAIATAFVLAVTWCGNEYEWASPMIIGLFIITLVAA 279
; QY 223 -----LTLINTGYFIP-YLHLVAHLQDL 245
; Db 280 IVFVFEKRAVDPLVPMGLFSNNFVLTAVAGIGVGLFMNGTIAVPTYLQMVHGLNPQT 339
; QY 246 WDPDLAFLLSVAISDLVGRVS-----GWLGDVAVPGPVTRLMLMTTLTGVSAL 297
; Db 340 AGMLIIPMIGLIGTSTVGNIVSKTGKYYKWPFIGMLI---MVALVLLSTLT----- 390
; QY 298 FPVAQAPTALVALAVAGFTSGALAPLAFSVLP-----ELIGTRRIYCGLQMI 348
; Db 391 -----PSASIALIGLYFVEFGGLGCAMQILVLIVQNSPITMTGT-----ATGSNNFF 439
; QY 349 ESIGLLGPLSG--YLRDVSQNT-----ASFVVAAGAFLLSG---SGILTLPHFF 395
; Db 440 RQIGAVGSAALIGLFIISNLSDRFTENPAVAVASMGEEGAQYASAMSDFGASNLTPHL- 498
; QY 396 CFSTTSGPQDL 407
; Db 499 ---VESLPQAL 506

```

RESULT 10
US-11-296-657-3
; Sequence 3, Application US/11296657
; Publication No. US20060143736A1
; GENERAL INFORMATION:
; APPLICANT: Schneeberger, Richard, Margolles-Clark, Emilio, Park, Joon-Hyun,
; TITLE OF INVENTION: MODULATING PLANT CARBON LEVELS
; FILE REFERENCE: 11696-149001
; CURRENT APPLICATION NUMBER: US/11/296,657
; CURRENT FILING DATE: 2005-12-06
; NUMBER OF SEQ ID NOS: 100
; SEQ ID NO 3

[illegible][illegible]


```

Db      213 AEKQVHESKAAYASATICALLFVPLTVSVKQELEVMNMKLP IEEPSEVKVEKKEL 272
QY      193 SLAEDFA--VGGPRAQLTSLH--HGPFLLR---YVALTLINTGYFIPYHLVAHL---- 241
Db      273 DLQDQKAAKVNGEKEKETKSCFSTVFSPPREDYTILOALLSTDMIILFVATFCGLGSSL 332
QY      242 -----QDLDMWDLPAAFLLSVVAISDLVGRVSVGWLGD-----AVEGPVTRLIM 285
Db      333 TAVDNLGQIGESLGYPNHTVSSFVSLVSIWNYFGRVFSGFVSEYLLAKYKLLPRPLMMTLV 392
QY      286 LWTTLTGVSLALFPVAQAPTALVALAVAGFTSGALAPLAFSVLPPELIGTR----RIYCG 341
Db      393 LLLSACGHLLIAFPV---PGSVYIASILMGFSFGAQLPLFAIISSELFGLKYYSTLFGC- 448
QY      342 LGLLQMIESTIGLLGPPLSGYLDRV 366
Db      449 -----GQLASPLGSIILNV 462

```

```

RESULT 13
US-11-056-355B-83753
; Sequence 83753, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2.
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966.
; SEQ ID NO 83753
; LENGTH: 577
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(577)
; OTHER INFORMATION: Ceres Seq. ID no. 12671663
; US-11-056-355B-83753

```

Query Match	5.9%	Score 138;	DB 7;	Length 577;
Best Local Similarity	22.1%	Pred. No. 0.0062;		
Matches 85;	Conservative 42;	Mismatches 114;	Indels 144;	Gaps 16;

	106	117	164	172	185	232	193	252	242	352	286	412
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L
b	L	L	S	H	Y	A	S	L	Y	L	L	L
Y	L	L	S	H	Y	A	S	L	Y	L	L	L

```

QY      342 LGLQMIESIGLLGPPLSGYLRDV 366
          | | | | : |
Db      468 -----GQLASPLGSYILNV 481

```

```

RESULT 14
US-11-330-403-1885
; Sequence 1885, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 1885
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Lactobacillus plantarum WCFS1
US-11-330-403-1885

```

Query Match	5.9%;	Score 137.5;	DB 7;	Length 470;
Best Local Similarity	22.4%;	Pred. NO. 0.0054;		
Matches	93;	Conservative 61;	Mismatches 158;	Indels 103; Gaps 19;
QY	22	FQSAVFEGVIRSFGEVFEVFAAFEEQARVSWIASIGIAV-QQFSGPVGSALSTKFGPR	80	
Db	34	YDTGIVNGSLFEMAVKGQLDLTAFQQ-----GIVSSGLTLGAFAIGAIIGGPFADKIGRK	87	
QY	81	PVVMGTGIIAALGMLLASFATSLTHYL---SIGLLSGS-----WALTFAF-----	124	
Db	88	KILTLIGLIFSVGALGCAFATNITLIVERFILGLAVGSASANVPVYIAEIAPTELRGKM	147	
QY	125	-TLACLSGYFSRRSLATGALTGVGLSSFTFAPFFQWLHSHYAWRGSLLVLSALSLHLV	183	
Db	148	VTTAQVMIVSGQFVAFGVNAAITPLGAQN---AAIWRMLGLTIPGIITLMI---GMYLIT	201	
QY	184	-----ACGALLRPPSLAE-----DPAVGGPRAQLTSLLHHG	214	
Db	202	PESPRMLVSGKMDKALGVLRIRRSASAVESEMEKEIQDKADKELNAAEQATFKELISK-	260	
QY	215	PFLRYTVALTLINTGYFIPYLHLVAH-----LQDLWDPLPAFL-----LSV	257	
Db	261	---RMVQILIT--TGAMLGIIQQFAGINSIMYGGKIIQESGFDTTVAAILNAGNPFISI	315	
QY	258	VAISDLVGRVSVGMLGDVAVEGPVTRLLMLWTTLTGVSLALFPVAQ--APTA-----LVA	309	
Db	316	VGA--VLGMFTIDWLGR-----RKLEFAGLTICGITLLVSAGVIHTVAPNASWAGITIVV	367	
QY	310	LAAVAY-GFTSGALAPLAFSVLPELIGTRRIYCGLLQMIESIG---GLIGPPL	359	
Db	368	LVYLYIIFFQGTIGPVTWMLINSEIFPQRYRGITGCTITPVLWIGNFIVGLSPVL	422	

```

RESULT 15
US-11-330-403-8385
; Sequence 8385, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 8385
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Lactobacillus brevis
US-11-330-403-8385

```

Query Match	5.9%;	Score 136.5;	DB 7;	Length 457;
-------------	-------	--------------	-------	-------------

Best Local Similarity 21.0%; Pred.No.0.0062;	
Matches 98; Conservative 82; Mismatches 156; Indels 131; Gaps 24;	
y	26 LVFGV---LRSFGVFVEFVAFAFEQAAKRVSWIASIGIAVQFGSPVGSAL----STKEG 78
b	18 LLFGYDTGVISGAILFI-----QKQNMIGSWQCGWVSALLGAILGAAIIGPSSDRFG 71
y	79 PRPVVMTGTLAALGMLASFATSLTHLYS---IGLLSGGWAAL--TEAPTLACLSCYF 133
b	72 RRKLLLSAIIFFVGALGSASFSPFWTLISRIILGMVAGAAALIPYLAELAP----- 126
y	134 SRRSLATGL----ALTGVGL---SSFTFAPFF---QMLSHYAMRGSLLVLSALSLHLV 183
b	127 SDRGTVSSSLFQLMWMTGILLAYITNYSFGFYTGWRMMLGFAAIPALLFL----- 178
y	184 ACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFLRYTVALTLINTGVFIPLYLHLVAHLQD 243
b	179 --GGLILPES-----PR-----FLVKSGHDEARHVLDTMNK 208
y	244 LDWDPLPAFLSLVVAISDLVGRVSGWLGDAVGPVTRLLMLMTTLTGVSIALLP---- 299
b	209 HD---QVAVNKEINDIQESAKIVSGW--SELFQKWRP---SLITIGLAIPOQVMG 258
y	300 ---VAQAPTLALVALVAYGFTSGALAPLAFSV-----LPELIGTRRIY----C 340
b	259 CNTVLYYAPT--IFTDVGFVSALLAHIGIIFNVIVTAIAVAIMDKIDRKKIWNIGAV 316
y	341 GIGLLQMIIESIGLIGPPLSGYLKRDVSGNYTASFVAVAGAFLLSGSGLLTLPHFFCFSTT 400
b	317 GMGISLFWMSIG-----MKFSGGSQTAAIT-----SVIALTV--YIAFFSA 355
y	401 TSGPQDLVTEALDTKVPPLPKEGLEGLNSTESGPESQ--SLTAPGLL 445
b	356 TWGP--VMVMVIGEVPFLNIRGLGNSFASVINWTFANMIVSLTFPSLL 400

Search completed: September 7, 2006, 13:12:46
Job time : 37 secs

Gencore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2006, 12:58:21 ; Search time 196 Seconds
(without alignments)
1066.061 Million cell updates/sec

Title: US-10-654-428-2
Perfect score: 2331
Sequence: 1 WARRTEPPDGGWGRVVVLSA.....SLTAPGILLPRLGLHRTTVP 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_8:*

1:	geneseqp1980s:*
2:	geneseqp1990s:*
3:	geneseqp2000s:*
4:	geneseqp2001s:*
5:	geneseqp2002s:*
6:	geneseqp2003as:*
7:	geneseqp2003bs:*
8:	geneseqp2004s:*
9:	geneseqp2005s:*
10:	geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2331	100.0	457	5	ABG30739 Abg30739 Human mon
2	2323	99.7	457	5	ABP69801 Abp69801 Human pol
3	2323	99.7	458	5	ABB80596 Abb80596 Human sbg
4	2154	92.4	426	5	ABP61019 Abp61019 Novel hum
5	2151	92.3	426	5	AAE16779 Aae16779 Human tra
6	2151	92.3	426	5	ABG32785 Abg32785 Human mon
7	2150	92.2	426	5	AD116770 Ad116770 Human NOV
8	2150	92.2	426	10	AEGL3486 Aegl3486 Amino aci
9	2143	91.9	426	8	ADN42424 Adn42424 Human nov
10	2142.5	91.9	427	5	AD116802 Ad116802 Human NOV
11	2142.5	91.9	427	5	AD117243 Ad117243 Human NOV
12	2142	91.9	426	5	AD116766 Ad116766 Human NOV
13	2142	91.9	426	5	AD116768 Ad116768 Human NOV
14	2142	91.9	426	8	ADN42420 Adn42420 Human nov
15	2142	91.9	426	8	ADN42422 Adn42422 Human nov
16	2142	91.9	426	10	AEGL3482 Aegl3482 Amino aci
17	2142	91.9	426	10	AEGL3484 Aegl3484 Amino aci
18	2142	91.9	426	10	AEGL3488 Aegl3488 Amino aci
19	941	40.4	186	4	AAU01618 Aau01618 Human sec
20	930.5	39.9	447	5	ABB80595 Abb80595 Human sbg
21	930.5	39.9	447	5	AD116488 Ad116488 Human NOV
22	930.5	39.9	471	4	AAW93737 Aaw93737 Human pol
23	930.5	39.9	471	8	AD131672 Ad131672 Human pro

24	920	39.5	472	5	AD116494 Ad116494 Human NOV
25	911	39.1	472	8	ADN42150 Adn42150 Human nov
26	900.5	38.6	447	8	ADN42144 Adn42144 Human nov
27	704	30.2	142	5	ADG79418 Adg79418 Human sec
28	704	30.2	142	5	ADG79508 Adg79508 Human sec
29	643	27.6	290	5	AD117855 Ad117855 Human NOV
30	640	27.5	486	6	AAO27118 Aao27118 Monocarob
31	637.5	27.3	486	5	AAE22711 Aae22711 Human tra
32	637.5	27.3	486	6	AAO27116 Aao27116 Monocarob
33	637.5	27.3	516	5	AAE22913 Aae22913 Human tra
34	637.5	27.3	516	6	AAO27117 Aao27117 Monocarob
35	637.5	27.3	516	7	ADG84149 Adg84149 Human TMD
36	634.5	27.2	486	6	AAO27119 Aao27119 Monocarob
37	634.5	27.2	486	8	ADR09111 Adr09111 Human pro
38	630	27.0	308	5	ABB80594 Abb80594 Human sbg
39	630	27.0	308	5	AAE21157 Aae21157 Human TRI
40	625.5	26.8	504	8	ADN29577 Adn29577 Human mon
41	625.5	26.8	504	9	AEA35823 Aea35823 Human mon
42	625.5	26.8	513	8	ADN29613 Adn29613 Strep-tag
43	625.5	26.8	513	9	AEA35859 Aea35859 Human mon
44	619.5	26.6	375	4	AAB88570 Aab88570 Human hyd
45	617	26.5	489	5	AD116804 Ad116804 Rat NOVX

ALIGNMENTS

RESULT 1	ABG30739	standard; protein; 457 AA.
ID	ABG30739	
XX	AC	ABG30739;
XX	DT	21-OCT-2002 (first entry)
XX	DE	Human monocarboxylate transporter related polypeptide.
XX	KM	Human; transporter; gene therapy; monocarboxylate transporter.
XX	OS	Homo sapiens.
XX	PN	WO200248364-A2.
XX	PD	20-JUN-2002.
XX	PF	11-DEC-2001; 2001WO-US047557.
XX	PR	12-DEC-2000; 2000US-0254554P.
PR	23-MAR-2001; 2001US-00815301.	
PR	10-APR-2001; 2001US-00829432.	
XX	PA	(PEKE) PE CORP NY.
XX	PI	Ketchum KA, Webster M, Merkulov G, Di Francesco V, Beasley EM;
XX	DR	WPI; 2002-583516/62.
XX	DR	N-PSDB; ABK89371.
PT	PT	Transporter proteins related to the monocarboxylate transporter
PT	PT	subfamily, for diagnosing and treating diseases mediated by the
PT	PT	transporter protein and identifying modulators.
XX	PS	Claim 1; Fig 2A; 84pp; English.
XX	CC	The invention relates to a human transporter protein and the
CC	CC	polynucleotide encoding it. The protein is useful for identifying a
CC	CC	modulator of the expression of the peptide by contacting the peptide with
CC	CC	an agent and determining if the agent has modulated the function or
CC	CC	activity or expression of the peptide. The sequences are useful as models
CC	CC	for the development of human therapeutic targets, aid in the
CC	CC	identification of therapeutic proteins and serve as targets for the
CC	CC	development of human therapeutic agents that modulate transporter
CC	CC	activity. The transporter proteins also provide a target for diagnosing a

disease or predisposition to a disease mediated by the protein, in pharmacogenomic analysis and for treating disorders characterised by an absence of inappropriate or unwanted expression of the protein. This sequence represents a human transporter protein, related to the monocarboxylate transporter subfamily

2 Sequence 457 AA;

Query Match 100.0%; Score 2331; DB 5; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.4e-205;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGI 60
b 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGI 60
Y 61 AVQOFGSPVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYSLIGLLSGSGMAL 120
b 61 AVQOFGSPVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYSLIGLLSGSGMAL 120
Y 121 TFAPTLACLSCYFSRRSLATGALTGVGSSFTFAPFQWLLSHYAWRGSLLVASLSL 180
b 121 TFAPTLACLSCYFSRRSLATGALTGVGSSFTFAPFQWLLSHYAWRGSLLVASLSL 180
Y 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPLYHLVAH 240
b 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPLYHLVAH 240
Y 241 LQDLWDPLPAFLLSVAISDLVGRVSGMLGDAVPGPTRLMLMTTLTGVSLLFPV 300
b 241 LQDLWDPLPAFLLSVAISDLVGRVSGMLGDAVPGPTRLMLMTTLTGVSLLFPV 300
Y 301 AQAPTALVALAVAGFTSGALAPLAFSVLPFLIGTRRIYCGLLQMIESIGLLGPPLS 360
b 301 AQAPTALVALAVAGFTSGALAPLAFSVLPFLIGTRRIYCGLLQMIESIGLLGPPLS 360
Y 361 GYL RDVSGNYTASFVVAAGAFLLSGS GILTLPHFECFSTTSGPQDLVTEALDTKVPLPK 420
b 361 GYL RDVSGNYTASFVVAAGAFLLSGS GILTLPHFECFSTTSGPQDLVTEALDTKVPLPK 420
Y 421 EGLEGGLNSTESGPESQSLTAPGILLPRLGLHRTTVP 457
b 421 EGLEGGLNSTESGPESQSLTAPGILLPRLGLHRTTVP 457

RESULT 2
BP69801
D ABP69801 standard; protein; 457 AA.

C ABP69801;

IT 20-JAN-2003 (first entry)

X Human polypeptide SEQ ID NO 1848.

X Human; genome mapping; gene therapy; food supplement; virus; fungus;
W cell-proliferative disorder; neurodegenerative disease; bacterial;
W Parkinson's disease; Alzheimer's disease; autoimmune disease;
W multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
W arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
W antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
W haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
W antiarthritic.

X Homo sapiens.

X WO200270539-A2.

X 12-SEP-2002.

X 05-MAR-2002; 2002WO-US005095.

X 05-MAR-2001; 2001US-00799451.

XX (HYSE-) HYSEQ INC.
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX MPI; 2002-759812/82.
DR N-PSDB; ABZ12018.

New polynucleotides comprising sequences assembled from expressed sequence tags (ESTs), useful for treating cell-proliferative, neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet or coagulation disorders.

Claim 9; SEQ ID NO 1848; 1012bp + Sequence Listing; English.

The invention relates to an isolated polynucleotide (I) comprising a nucleotide sequence selected from any of 948 sequences (ABZ1119-ABZ12066) or their mature protein coding portion, active domain coding protein or complementary sequences. The polynucleotides are useful for identifying expressed genes or for physical mapping of human genome. The encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight markers, as a food supplement, for generating antibodies, in medical imaging, screening and diagnostic assays and for treating cell-proliferative disorders (cancer), neurodegenerative diseases (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, platelet or coagulation disorders, wound, burns, incision, ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal, parasitic), arthritis, etc. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 457 AA;

Query Match 99.7%; Score 2323; DB 5; Length 457;
Best Local Similarity 99.8%; Pred. No. 7.8e-205;
Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGI 60
Db 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGI 60
QY 61 AVQOFGSPVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYSLIGLLSGSGMAL 120
Db 61 AVQOFGSPVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYSLIGLLSGSGMAL 120
QY 121 TFAPTLACLSCYFSRRSLATGALTGVGSSFTFAPFQWLLSHYAWRGSLLVASLSL 180
Db 121 TFAPTLACLSCYFSRRSLATGALTGVGSSFTFAPFQWLLSHYAWRGSLLVASLSL 180
QY 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPLYHLVAH 240
Db 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPLYHLVAH 240
QY 241 LQDLWDPLPAFLLSVAISDLVGRVSGMLGDAVPGPTRLMLMTTLTGVSLLFPV 300
Db 241 LQDLWDPLPAFLLSVAISDLVGRVSGMLGDAVPGPTRLMLMTTLTGVSLLFPV 300
QY 301 AQAPTALVALAVAGFTSGALAPLAFSVLPFLIGTRRIYCGLLQMIESIGLLGPPLS 360
Db 301 AQAPTALVALAVAGFTSGALAPLAFSVLPFLIGTRRIYCGLLQMIESIGLLGPPLS 360
QY 361 GYL RDVSGNYTASFVVAAGAFLLSGS GILTLPHFECFSTTSGPQDLVTEALDTKVPLPK 420
Db 361 GYL RDVSGNYTASFVVAAGAFLLSGS GILTLPHFECFSTTSGPQDLVTEALDTKVPLPK 420
QY 421 EGLEGGLNSTESGPESQSLTAPGILLPRLGLHRTTVP 457
Db 421 EGLEGGLNSTESGPESQSLTAPGILLPRLGLHRTTVP 457

```
RESULT 3
ABB80596
ID ABB80596 standard; protein; 458 AA.
XX AC ABB80596;
XX 08-OCT-2002 (first entry)
DT
XX Human sbg102200MCTb protein.
DE
XX
KW Human; secreted protein; immunosuppressive; cytostatic; nootropic;
KW neuroprotective; antitumour; vulnery; antimicrobial; ophthalmological;
KW antiparkinsonian; antirheumatic; atherosclerotic; dermatological;
KW hypotensive; cerebroprotective; virucide; antiinflammatory; diabetes;
KW malignant tumour; hypertension; hypotension; obesity; bulimia; anorexia;
KW asthma; manic depression; Tourette's syndrome; schizophrenia;
KW Huntington's disease; dementia; delirium; mental retardation;
KW mental disorder; sexual development disorder; blood cascade dysfunction;
KW stroke; growth disorder.
XX
OS Homo sapiens.
XX WO200222802-A1.
PN
XX 21-MAR-2002.
PD
XX
PF 13-SEP-2001; 2001WO-US028462.
XX
PR 13-SEP-2000; 2000US-0232455P.
PR 13-SEP-2000; 2000US-0232463P.
PR 02-OCT-2000; 2000US-0237293P.
PR 07-NOV-2000; 2000US-0246269P.
PR 20-NOV-2000; 2000US-0252049P.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Agarwal P, Gogswell JP, Lai Y, Martensen SA, Smith RF, Strum JC,
PI Xie Q;
XX
DR WPI; 2002-393963/42.
DR N-PSDB; ABL57732.
XX
XX Novel isolated secreted polypeptides and polynucleotides encoding them
PT useful for treating cancer, Alzheimer's disease, tumor metastasis,
PT autosomal recessive atypical hemolytic uremic syndrome, wound healing
PT disorder.
XX
PS 1; Page216-217; 246pp; English.
XX
CC The invention relates to a novel isolated polypeptide (ABB80569-ABB80612)
CC (secreted polypeptide) which is encoded by any one of 44 polynucleotide
CC sequences (AB157705-ABL57748) given in the specification. The
CC polypeptides have immunosuppressive, cytostatic, nootropic,
CC neuroprotective, antitumour, vulnery, antineurobiol, ophthalmological,
CC antiparkinsonian, antirheumatic, atherosclerotic, dermatological,
CC hypotensive, cerebroprotective, virucide, and antiinflammatory activity.
CC The polynucleotide and polypeptide are useful for treating diabetes,
CC malignant tumours, hyper- and hypotension, obesity, bulimia, anorexia,
CC asthma, manic depression, dementia, delirium, mental retardation,
CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental
CC or sexual development disorders, and dysfunctions of the blood cascade
CC system including those leading to stroke
XX
XX Sequence 458 AA;
30
```

```
Query Match 99.7%; Score 2323; DB 5; Length 458;
Best Local Similarity 99.8%; Pred. No. 7.8e-205;
Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
1 MARRTEPPDGGWGVVLSAFQSAALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGI 60
|||||
```

```
Db 2 MARRTEPPDGGWGVVLSAFQSAALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGI 61
QY AVOQFGSPVGSALSTKFGRRPVMTGGILALGMLASFATSLTHLYLSIGLSSGSAW 120
Db 62 AVOQFGSPVGSALSTKFGRRPVMTGGILALGMLASFATSLTHLYLSIGLSSGSAW 121
QY 121 TFAPTLACLSCFSRRSLATGLALTGVGLSFTPAFFQMLSHYAWRGSLLVASL 180
Db 122 TFAPTLACLSCFSRRSLATGLALTGVGLSFTPAFFQMLSHYAWRGSLLVASL 181
QY 181 HLVACGALLRPPSLAEDPAVGPPRAQLTSLHGPFLRYTVALTLINTGYFIPYLHVAH 240
Db 182 HLVACGALLRPPSLAEDPAVGPPRAQLTSLHGPFLRYTVALTLINTGYFIPYLHVAH 241
QY 241 LQDLWDPLPAFLISVVAISDLVGRVVSGLGDAVPGPVTRLMLWTTLTGVSIALFPV 300
Db 242 LQDLWDPLPAFLISVVAISDLVGRVVSGLGDAVPGPVTRLMLWTTLTGVSIALFPV 301
QY 301 AQAPTALVALAVAGFTSGALAPLAFSVLPFLIGTRRIYCGLLQMIIESIGLIGPPLS 360
Db 302 AQAPTALVALAVAGFTSGALAPLAFSVLPFLIGTRRIYCGLLQMIIESIGLIGPPLS 361
QY 361 GYLKDVSGNYTASFVAGAPFLISGSGILTLPHFFCFSTTSGPQDLVTEALDTKVPPLPK 420
Db 362 GYLKDVSGNYTASFVAGAPFLISGSGILTLPHFFCFSTTSGPQDLVTEALDTKVPPLPK 421
QY 421 EGLEGGLNSTESGPESQSLTAPGLLPRLGHRHTVP 457
Db 422 EGLEGGLNSTESGPESQSLTAPGLLPRLGHRHTVP 458
```

```
RESULT 4
ABP61019
ID ABP61019 standard; protein; 426 AA.
XX
AC ABP61019;
XX
DT 10-SEP-2002 (first entry)
XX
DE Novel human protein. SEQ ID 110.
XX
KW Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW nootropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antilucer; virucide; antithyroid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty.
XX
OS Homo sapiens.
XX
PN WO200250105-A1.
XX
PD 27-JUN-2002.
XX
PF 17-DEC-2001; 2001WO-US049232.
XX
PR 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
```

PR 08-MAY-2001; 2001US-0289622P.
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
DX WPI; 2002-508784/54.
DR N-PSDB; AB086184.
XX
PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
PS Claim 1(a); Page 331-332; 335pp; English.
XX
CC The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytosstatic, vulnerary, antiarteriosclerotic, antiparkinsonian, nootropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, antilicer, virucide, antithyroid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, angiogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records ABP60965-ABP61019 represent novel human
CC proteins of the invention
CC
XX
SQ Sequence 426 AA;
Query Match 92.4%; Score 2154; DB 5; Length 426;
Best Local Similarity 99.8%; Pred. No. 2.6e-189;
Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 361 GYLKRVSGNNTASFWVAGAFLLSGSGILLTLPHFCEFSTTSGPDVLTEALDTKVLPLK 420
QY 421 EGLE 424
Db 421 EGLE 424
RESULT 5
AAE16779
ID AAE16779 standard; protein; 426 AA.
XX
AC AAE16779;
XX
XX 09-APR-2002 (first entry)
DT
XX
DE Human transporter and ion channel-16 (TRICH-16) protein.
XX
KW Human; transporter and ion channel-16; TRICH-16; neuroprotective; asthma;
KW nootropic; cytosstatic; cardiovascular; immunosuppressive; cardiomyopathy;
KW antiinflammatory; protein therapy; akinesia; cystic fibrosis; leukaemia;
KW Bell's palsy; amyotrophic lateral sclerosis; Alzheimer's disease; cancer;
KW amnesia; dementia; myocarditis; Duchenne's muscular dystrophy; AIDS;
KW Acquired Immune Deficiency Syndrome; Addison's disease; allergy; angina;
KW cell proliferative disorder; psoriasis; cardiac disease; hypertension;
KW bradycardia; gene expression; drug screening.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Domain 82..109
FT /label= Transmembrane_domain
FT Domain 338..356
FT /label= Transmembrane_domain
XX
PN WO200192304-A2.
XX
XX
PD 06-DEC-2001.
PD
XX
PF 25-MAY-2001; 2001WO-US017065.
PF
XX
PR 26-MAY-2000; 2000US-0208424P.
PR 01-JUN-2000; 2000US-0209001P.
PR 08-JUN-2000; 2000US-0210588P.
PR 16-JUN-2000; 2000US-0212335P.
PR 22-JUN-2000; 2000US-0213747P.
PR 29-JUN-2000; 2000US-0215391P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Thornton M, Walia NK, Yue H, Nguyen DB, Lai P, Gandhi AR;
PI Tribouley CM, Yao MG, Ramkumar J, Au-Young J, Lu Y, Tang YT;
PI Azimzai Y, Bruns CM, Griffin JA, Yang J, Sanjanwala MS, Raumann BE;
PI Lee EA, Hafalia A, Baughn MR, Green BD, Khan FA, Kearney L;
PI Elliot VS, Seilhamer JJ, Policky JL, Borowsky ML, Burford N, Ding L;
PI Lu DAM, Hillman JL;
XX
XX WPI; 2002-122055/16.
DR
DR N-PSDB; AAD27269.
XX
PT New human transporters and ion channels (TRICH) polypeptides useful for
PT diagnosing, treating or preventing disorders associated with aberrant
PT expression of TRICH.
XX
PS Claim 1; Page 166; 210pp; English.
XX
XX The invention relates to human transporters and ion channels (TRICH)
CC polypeptides and their cDNA molecules. The nucleic acid and polypeptide
CC sequences are useful in the diagnosis, treatment, and prevention of
CC disorders associated with transport (akinesia, cystic fibrosis, Bell's
CC palsy, amyotrophic lateral sclerosis); neurological (Alzheimer's disease,
CC amnesia, dementia); muscle (cardiomyopathy, myocarditis, Duchenne's
CC muscular dystrophy); immunological (AIDS, Addison's disease, allergies,
CC asthma); cell proliferative disorders (cancers, leukaemia, psoriasis);

CC cardiac disease (angina, hypertension, or bradyarrhythmia) and in the
CC assessment of the effects of exogenous compounds on the expression of
CC nucleic acid and amino acid sequences of transporters and ion channels.
CC The polynucleotides may be used to detect and quantify gene expression in
CC biopsied tissues in which TRICH expression may be correlated with a
CC disease, to generate hybridization probes for mapping naturally occurring
CC genomic sequence, and in drug screening. The present sequence is human
CC TRICH-16 protein
XX
SQ Sequence 426 AA;

Query Match 92.3%; Score 2151; DB 5; Length 426;
Best Local Similarity 99.5%; Pred. No. 4.9e-189;
Matches 422; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	MARRTEPPDGGWGRVVVLSAFQSA	LVFGVLRSGVFVEFVA	AEEQAA	RVSWIASIGI	60
Db	1	MARRTEPPDGGWGVVLSAFQSA	LVFGVLRSGVFVEFVA	AEEQAA	RVSWIASIGI	60
QY	61	AVQFGSPVGSALSTKFGPRPVMTG	GILALGMLLASFATSLTHLYS	IGLLSGSGWAL	120	
Db	61	AVQFGSPVGSALSTKFGPRPVMTG	GILALGMLLASFATSLTHLYS	IGLLSGSGWAL	120	
QY	121	TFAPTLACLSYFSRRSLATGALT	GVGLSSFTFAPFQWLLSHYAW	RGSLLVSALSL	180	
Db	121	TFAPTLACLSYFSRRSLATGALT	GVGLSSFTFAPFQWLLSHYAW	RGSLLVSALSL	180	
QY	181	HLVACGALLRPPSLAEDPAVG	GPRAQLTSLHHGPFLRYTVAL	TLINTGYFIPLYHLVAH	240	
Db	181	HLVACGALLRPPSLAEDPAVG	GPRAQLTSLHHGPFLRYTVAL	TLINTGYFIPLYHLVAH	240	
QY	241	LQDLWDPLPAAFLLSVVAISD	LVGRVSGWLGDAVPGVTRLML	MTTLTGVSALFPV	300	
Db	241	LQDLWDPLPAAFLLSVVAISD	LVGRVSGWLGDAVPGVTRLML	MTTLTGVSALFPV	300	
QY	301	AOAPTALVALAVAYGFTSGAL	APLAFSVLPBELIGTRRIYCG	LGLQMI	ESIGLLGPPLS	360
Db	301	AOAPTALVALAVAYGFTSGAL	APLAFSVLPBELIGTRRIYCG	LGLQMI	ESIGLLGPPLS	360
QY	361	GYLRDVSGNYTASFVVA	GAFLLSGS	GILLTLPHF	CFSTTSGPQDLVTEAL	DTKVPLPK 420
Db	361	GYLRDVSGNYTASFVVA	GAFLLSGS	GILLTLPHF	CFSTTSGPQDLVTEAL	DTKVPLPK 420
QY	421	EGLE	424			
Db	421	EGLE	424			

RESULT 6
ABG32785
ID ABG32785 standard; protein; 426 AA.
XX
AC ABG32785;

DT 29-NOV-2002 (first entry)

DE Human monocarboxylic acid transporter protein TCH081 protein.
XX
KW Human; hepatotropic; antidiabetic; nephrotropic; immunostimulant;
KW immunosuppressant; antiinflammatory; antiallergic; cytostatic; antitumor;
KW vasotropic; neuroprotective; nootropic; liver disease; diabetes;
KW kidney disease; metabolic acidosis; muscular disease; spleen disease;
KW immunodeficiency; autoimmune disease; inflammatory disease; allergy;
KW genital disease; digestive disease; central nervous system disorder;
KW circulatory disease; cancer;
KW monocarboxylic acid transporter protein TCH081.

XX Homo sapiens.
XX
XX WO200262998-A1.
XX
XX 15-AUG-2002.
XX

PF 05-FEB-2002; 2002WO-JP000914.
XX
XX 06-FEB-2001; 2001JP-00030172.
PR 21-JUN-2001; 2001JP-00188708.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Nakanishi A, Segiya Y;
XX
XX WPI; 2002-627556/67.
DR N-PSDB; ABS53751.
XX
PT Human monocarboxylic acid transporter protein TCH081 and DNA encoding it
PT for screening compounds modifying its activity which can be used as
PT drugs.
XX
XX Claim 1; Page 82-84; 102pp; Japanese.

CC This invention relates to the DNA and protein sequences of a human
CC monocarboxylic acid transporter protein TCH081 and its salts. The protein
CC of the invention may have hepatotropic, antidiabetic, nephrotropic,
CC immunostimulant, immunosuppressant, antiinflammatory, antiallergic,
CC cyostatic, antitumor, vasotropic, neuroprotective and nootropic
CC activity. Modification of transport of monocarboxylic acids and amino
CC acids into cells mediated by the TCH081 protein. The protein of the
CC invention is useful for the treatment, prevention and diagnosis of a wide
CC range of disorders including liver disease, diabetes, kidney disease,
CC metabolic acidosis, muscular disease, spleen disease, immunodeficiency,
CC autoimmune diseases, inflammatory diseases, allergies, genital diseases,
CC digestive diseases and cancer. The present sequence represents the human
CC monocarboxylic acid transporter protein TCH081 of the invention
XX
SQ Sequence 426 AA;

Query Match 92.3%; Score 2151; DB 5; Length 426;
Best Local Similarity 99.5%; Pred. No. 4.9e-189;
Matches 422; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	MARRTEPPDGGWGRVVVLSAFQSA	LVFGVLRSGVFVEFVA	AEEQAA	RVSWIASIGI	60
Db	1	MARRTEPPDGGWGVVLSAFQSA	LVFGVLRSGVFVEFVA	AEEQAA	RVSWIASIGI	60
QY	61	AVQFGSPVGSALSTKFGPRPVMTG	GILALGMLLASFATSLTHLYS	IGLLSGSGWAL	120	
Db	61	AVQFGSPVGSALSTKFGPRPVMTG	GILALGMLLASFATSLTHLYS	IGLLSGSGWAL	120	
QY	121	TFAPTLACLSYFSRRSLATGALT	GVGLSSFTFAPFQWLLSHYAW	RGSLLVSALSL	180	
Db	121	TFAPTLACLSYFSRRSLATGALT	GVGLSSFTFAPFQWLLSHYAW	RGSLLVSALSL	180	
QY	181	HLVACGALLRPPSLAEDPAVG	GPRAQLTSLHHGPFLRYTVAL	TLINTGYFIPLYHLVAH	240	
Db	181	HLVACGALLRPPSLAEDPAVG	GPRAQLTSLHHGPFLRYTVAL	TLINTGYFIPLYHLVAH	240	
QY	241	LQDLWDPLPAAFLLSVVAISD	LVGRVSGWLGDAVPGVTRLML	MTTLTGVSALFPV	300	
Db	241	LQDLWDPLPAAFLLSVVAISD	LVGRVSGWLGDAVPGVTRLML	MTTLTGVSALFPV	300	
QY	301	AOAPTALVALAVAYGFTSGAL	APLAFSVLPBELIGTRRIYCG	LGLQMI	ESIGLLGPPLS	360
Db	301	AOAPTALVALAVAYGFTSGAL	APLAFSVLPBELIGTRRIYCG	LGLQMI	ESIGLLGPPLS	360
QY	361	GYLRDVSGNYTASFVVA	GAFLLSGS	GILLTLPHF	CFSTTSGPQDLVTEAL	DTKVPLPK 420
Db	361	GYLRDVSGNYTASFVVA	GAFLLSGS	GILLTLPHF	CFSTTSGPQDLVTEAL	DTKVPLPK 420
QY	421	EGLE	424			
Db	421	EGLE	424			

RESULT 7

ID116770
ID AD116770 standard; protein; 426 AA.
XC
XC AD116770;
XC
XC 15-APR-2004 (first entry)
XC
XC Human NOVX protein to treat human pathological conditions SeqID306.
XC
XC human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
XC inflammation; autoimmune disorder; allergy; blood disorder;
XC acquired immunodeficiency syndrome; AIDS; obesity; asthma;
XC immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
XC Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;
XC cytosstatic; cardiant; antiinflammatory; immunosuppressive; antiallergic;
XC haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; anorectic;
XC antiasthmatic; nephrotropic; antiarthritic; hepatotropic;
XC neuroprotective; nootropic; antibacterial; virucide; antiparasitic;
XC relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis;
XC chromosome mapping; tissue typing; pharmacogenomic.
XC
XC Homo sapiens.
XC
XC WO200268649-A2.
XC
XC 06-SEP-2002.
XC
XC 31-JAN-2002; 2002WO-US002785.
XC
XC 31-JAN-2001; 2001US-0265395P.
XC 31-JAN-2001; 2001US-0265412P.
XC 31-JAN-2001; 2001US-0265514P.
XC 31-JAN-2001; 2001US-0265517P.
XC 02-FEB-2001; 2001US-0266406P.
XC 05-FEB-2001; 2001US-0266767P.
XC 07-FEB-2001; 2001US-0266975P.
XC 07-FEB-2001; 2001US-0267057P.
XC 08-FEB-2001; 2001US-0267459P.
XC 09-FEB-2001; 2001US-0267823P.
XC 15-FEB-2001; 2001US-0268974P.
XC 26-FEB-2001; 2001US-0271664P.
XC 27-FEB-2001; 2001US-0271839P.
XC 27-FEB-2001; 2001US-0271855P.
XC 02-MAR-2001; 2001US-0272788P.
XC 02-MAR-2001; 2001US-0273046P.
XC 14-MAR-2001; 2001US-0275925P.
XC 14-MAR-2001; 2001US-0275947P.
XC 14-MAR-2001; 2001US-0275950P.
XC 14-MAR-2001; 2001US-0275989P.
XC 15-MAR-2001; 2001US-0276448P.
XC 15-MAR-2001; 2001US-0276450P.
XC 16-MAR-2001; 2001US-0276397P.
XC 16-MAR-2001; 2001US-0276768P.
XC 20-MAR-2001; 2001US-0278652P.
XC 26-MAR-2001; 2001US-0278775P.
XC 26-MAR-2001; 2001US-0278778P.
XC 29-MAR-2001; 2001US-0279882P.
XC 29-MAR-2001; 2001US-0279884P.
XC 30-MAR-2001; 2001US-0280147P.
XC 11-APR-2001; 2001US-0282992P.
XC 11-APR-2001; 2001US-0283083P.
XC 20-APR-2001; 2001US-0285133P.
XC 23-APR-2001; 2001US-0285749P.
XC 03-MAY-2001; 2001US-0288327P.
XC 03-MAY-2001; 2001US-0288504P.
XC 29-MAY-2001; 2001US-0294047P.
XC 30-MAY-2001; 2001US-0294473P.
XC 08-JUN-2001; 2001US-0296964P.
XC 18-JUN-2001; 2001US-0298959P.
XC 19-JUN-2001; 2001US-0299324P.
XC 13-AUG-2001; 2001US-0312020P.
XC 16-AUG-2001; 2001US-0312889P.
XC 16-AUG-2001; 2001US-0312908P.

PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XC
XC (CURA-) CURAGEN CORP.
XC
XC Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
XC Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
XC Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
XC Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XC
XC WPI; 2002-706998/76.
XC DR N-PSDB; AD116769.
XC
XC New NOVX polypeptides and nucleic acids, useful for preventing or
XC treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XC atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XC pharmacogenomics.
XC
XC Claim 1; SEQ ID NO 306; 1498pp; English.
XC
XC This invention relates to a novel nucleic acids, and encoded polypeptides
XC thereof, which have properties related to the stimulation of biochemical
XC or physiological responses in a cell, tissue, organ or organism.
XC Specifically, it refers to the use of biologically active fragments for
XC diagnostic and prognostic assays and furthermore in the treatment of
XC diverse pathological conditions. The present invention describes novel
XC human and murine NOVX proteins, as well as methods to modulate their
XC expression using antisense oligos, ribozymes and peptide nucleic acids.
XC The NOVX polypeptides, polynucleotides and antibodies are useful in
XC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
XC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
XC treating or preventing diseases such as inflammation, autoimmune
XC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
XC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
XC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
XC and epilepsy. Accordingly, these molecules have many activities including
XC cytosstatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
XC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
XC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
XC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
XC relaxant and anticonvulsant. In addition, they are useful in screening
XC assays to identify small molecules that modulate or inhibit, for example,
XC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
XC used as in chromosome mapping, tissue typing, preventive medicine and
XC pharmacogenomics. This polypeptide is a human NOVX protein of the
XC invention.
XC
XC Sequence 426 AA;
SQ
Query Match 92.2%; Score 2150; DB 5; length 426;
Best Local Similarity 99.5%; Pred. No. 6e-189;
Matches 422; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLRFGVFVEFVAAFEEOAARVSWIASIGI 60
Db 1 MARRTEPPDGGWGVVVVLSAFFQSALVFGVLRFGVFVEFVAAFEEOAARVSWIASIGI 60
QY 61 AVQQFGSPVGSALSTKFGPRPVMTGGILAAAGMLLASFATSLTHLYSIGLSSGSWAL 120
Db 61 AVQQFGSPVGSALSTKFGPRPVMTGGILAAAGMLLASFATSLTHLYSIGLSSGSWAL 120
QY 121 TFAPTLACLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLSHYAWRGSLLVASLSL 180
Db 121 TFAPSLACLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLSHYAWRGSLLVASLSL 180

QY	181	HLVACGALLRPPSLAEDPAVGGPRAOQLTSLHHGPFRLRYVALTLINTGYFIPYHLVAH	240
Db	181	HLVACGALLRPPSLAEDPAVGGPRAOQLTSLHHGPFRLRYVALTLINTGYFIPYHLVAH	240
QY	241	LQDLWDMDPLPAAFLLSVVAISDLVGRVVSQWLGDAVPGPVTRLMLMTTLTGVSALLFPV	300
Db	241	LQDLWDMDPLPAAFLLSVVAISDLVGRVVSQWLGDAVPGPVTRLMLMTTLTGVSALLFPV	300
QY	301	AQAPFTALVALAVAYGFTSGALAPLAFSVLPBELLIGTRRIYCGLLQMIESTIGLLGPPLS	360
Db	301	AQAPFTALVALAVAYGFTSGALAPLAFSVLPBELLIGTRRIYCGLLQMIESTIGLLGPPLS	360
QY	361	GYLRDVSQNYTASFVAVAGAFLLSGSGILLTLPHFFCFSTTSGPQDLVTEADTKVPLPK	420
Db	361	GYLRDVSQNYTASFVAVAGAFLLSGSGILLTLPHFFCFSTTSGPQDLVTEADTKVPLPK	420
QY	421	EGLE 424	
Db	421	EGLE 424	

```

RESULT 8
      AEG13486
ID      AEG13486 standard; protein; 426 AA.
XX
XX      AEG13486;
AC
XX
DT      20-APR-2006 (first entry)
XX
DE      Amino acid sequence for human NOV6c.

```

KW NOV polypeptide; substrate; screening; obesity; diabetes;
KW myocardial disease; atherosclerosis; hypertension; scleroderma;
KW prostate tumor; acquired immune deficiency syndrome; asthma;
KW multiple sclerosis; infectious disease; anorexia nervosa;
KW Alzheimers disease; parkinsons disease; anorectic; antidiabetic;
KW cardiant; cardiovascular-gen.; antiarteriosclerotic; hypotensive;
KW dermatological; cytostatic; anti-hiv; antiasthmatic; neuroprotective;
KW antimicrobial; anabolic; eating-disorders-gen.; nootropic;
KW antiparkinsonian.

Identifying compounds that modulate target polypeptide activity, by combining test compound with target polypeptide and substrate of the target polypeptide, and determining if the compound modulates activity of the target polypeptide.

The invention relates to a method of identifying compounds that modulate target polypeptide activity. The method comprises: (a) combining a test compound with a target polypeptide and a substrate of the target polypeptide; and (b) determining whether the test compound modulates the activity of the target polypeptide, where the target polypeptide comprises an amino acid sequence selected from SEQ ID NO. 2n, where n is 1-16, the amino acid sequence that is at least 95 % identical to SEQ ID

CC NO. 2n, the amino acid sequence of at least one domain of SEQ ID NO. 2n,
CC or the amino acid sequence that is at least 95 % identical to the at
CC least one domain of SEQ ID NO. 2n. Also described are: (1) an antibody
CC that immunospecifically binds to the target polypeptide; (2) a method of
CC identifying a potentially therapeutic agent for use in treatment of a
CC pathology, where the pathology is related to aberrant expression or
CC aberrant physiological interactions of a target polypeptide; (3)
CC screening for a modulator of activity or of latency or predisposition to
CC pathology associated with a target polypeptide; and (4) a nucleic acid
CC sequence encoding a target polypeptide of the invention, where the
CC nucleic acid comprises a sequence selected from SEQ ID NO. 2n-1. The
CC method of the invention is useful for identifying compounds that modulate
CC target polypeptide activity. The compounds can be used for the treatment
CC of obesity or diabetes. It can also be used for treating pathology
CC related to aberrant expression or aberrant physiological interactions of
CC a target polypeptide. Pathologies include cardiomyopathy,
CC atherosclerosis, hypertension, scleroderma, prostate cancer, AIDS,
CC bronchial asthma, multiple sclerosis, infectious disease, anorexia,
CC Alzheimer's disease, or Parkinson's disease. This sequence represents a
CC human NOV protein.
XX
SQ Sequence 426 AA;

Query Match	92.2%;	Score 2150;	DB 10;	Length 426;
Best Local Similarity	99.5%;	Pred. No. 6e-189;		
Matches 422; Conservative	1;	Mismatches 1;	Indels 0.	Gaps 0

QY 1 MARRTPEPDGCGWRVVLVSFAFFQSALVEGVLRSGVFVEFVAAFEQARVSWIASIGI 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MARRTEPDGCGMWVVLVSFAFFQSALVEGVLRSGVFVEFVAAFEQARVSWIASIGI 60

QY		61 AVQQFGSPVGSALSTKFGPRPVMTGGILAAIAGMLLASFATSLTHLYLSIGLLSGGMAI 120
Db	61	AVQQFGSPVGSALSTKFGPRPVMTGGILAAIAGMLLASFATSLTHLYLSIGLLSGGMAI 120

QY	121	TEAPFLACLS	CFSRRRS	LATG	LALTGV	GSSFTF	AFFQW	LISHY	AMRGS	LTLVSA	ISL	180
Dp	121	TEAPSLAC	SCFSRRRS	LATG	LALTGV	GSSFTF	AFFQW	LISHY	AMRGS	LTLVSA	ISL	180

QY	Db
181 HLVACGALLRPPSLAEDPAVGGRPQQLTSLHHGPRLRYTVALTLINTGYFIPYHLVAH 240	181 HLVACGALLRPPSLAEDPAVGGRPQQLTSLHHGPRLRYTVALTLINTGYFIPYHLVAH 240

QY	241	QDLDMDPLPAAFLLSVVAISDLVGRVSGWLGDVPGPVTRLMLWTTLLTGVSLALFPV	300
Dp	241	LQDLDMDPLPAAFLLSVVAISDLVGRVSGWLGDVPGPVTRLMLWTTLLTGVSLALFPV	300

QY	301	AQAPTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGGLQLQMIESIGLLGPPLS	360
Db	301	AQAPTALVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGGLQLQMIESIGLLGPPLS	360

QY	361	GYLRDVS	GN	Y	T	A	S	F	V	V	A	G	A	F	L	L	S	G	I	L	T	L	P	H	F	F	C	F	S	T	T	S	G	P	D	L	T	E	A	L	D	T	K	V	L	P	K	420
Dd	361	GYLRDVS	GN	Y	T	A	S	F	V	V	A	G	A	F	L	L	S	G	I	L	T	L	P	H	F	F	C	F	S	T	T	S	G	P	D	L	T	E	A	L	D	T	K	V	L	P	K	420

QY	4 21	EGLE	4 24
Db	4 21	EGLE	4 24

RESULT 9
ADN42424
ADN42424 standard; protein; 426 AA

17-JUN-2004 (first entry)
Human novel protein NOV 96c.

Human; NOVX; cancer; diabetes; cardiomyopathy; atherosclerosis.
Homo sapiens.

XX US2004033493-A1.
XX 19-FEB-2004.
XX 31-JAN-2002; 2002US-00072012.
PR 31-JAN-2001; 2001US-0265395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266767P.
PR 07-FEB-2001; 2001US-0266975P.
PR 07-FEB-2001; 2001US-0267057P.
PR 08-FEB-2001; 2001US-0267459P.
PR 09-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 02-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 14-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313930P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX
PA (TCHE/) TCHERNEV V T.
PA (SPYT/) SPYTEK K A.
PA (ZERH/) ZERHUSEN B D.
PA (PATI/) PATTURAJAN M.
PA (SHIM/) SHIMKETS R A.
PA (LILU/) LI L.
PA (GANG/) GANGOLLI E A.
PA (PADI/) PADIGARU M.
PA (ANDE/) ANDERSON D W.
PA (RAST/) RASTELLI L.
PA (MILL/) MILLER C E.

PA (GERL/) GERLACH V.
PA (TAUP/) TAUPIER R J.
PA (GUSE/) GUSEV V Y.
PA (COLM/) COLMAN S D.
PA (WOLE/) WOLENC A R.
PA (PENA/) PENNA C E A.
PA (FURT/) FURTAK K.
PA (GROS/) GROSSE W M.
PA (ALSO/) ALSOBROOK J P.
PA (LEPL/) LEPLEY D M.
PA (RIEG/) RIEGER D K.
PA (BURG/) BURGESS C E.
XX
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
PI Gerlach V, Taupier RU, Gusev VY, Colman SD, Wolenc AR, Pena CE;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
XX WPI; 2004-180039/17.
DR N-PSDB; ADN42423.
XX
XX Isolated NOVX polypeptides and polynucleotides, useful for preventing
PT diagnosing and/or treating cancer, diabetes, cardiomyopathy and
PT atherosclerosis.
XX
XX Claim 1; SEQ ID NO 306; 1309pp; English.
PS
XX
XX The invention relates isolated 162 NOVX polypeptides (NOV1-NOV99,
CC including splice variants) and the nucleic acids (NA) that encode them.
CC Also included are the mature NOVX proteins (and their encoding
CC polynucleotides), a vector comprising NOVX NA, a cell comprising the
CC vector, an antibody that binds immunospecifically to NOVX, determining
CC the presence or amount of NOVX in a sample, determining the presence or
CC amount of NOVX NA in a sample, identifying an agent that binds to NOVX,
CC modulating the activity of NOVX, treating or preventing a NOVX-associated
CC disorder, determining the presence of or predisposition to a disease
CC associated with altered levels of NOVX and treating a pathological state
CC in a mammal comprising administering a polypeptide which is at least 95%
CC identical to NOVX (or fragment). NOVX and NA may be used in the
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression and activity of NOVX (e.g. cancer, diabetes,
CC cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and
CC antagonists may also be used to down regulate expression and activity of
CC NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for
CC detecting the presence of NOVX in samples (e.g. by enzyme linked
CC immunosorbant assay (ELISA). The agents and methods may be used in this
CC way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy
CC and/or atherosclerosis. The present sequence represents a NOVX protein.
XX
XX Sequence 426 AA;
SQ
Query Match 91.9%; Score 2143; DB 8; Length 426;
Best Local Similarity 99.3%; Pred. No. 2.7e-188;
Matches 421; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MARRTEPPDGGWGRVVVLSA FQSA LVFGVRSFGVFVEFVA AFEEQAA RVSWIASIGI 60
Db 1 MARRTEPPDGGWGVVVLSA FQSA LVFGVRSFGVFVEFVA AFEEQAA RVSWIASIGI 60
QY 61 AVQQFGSPVGSALSTKFGPRPVVMTGGILALGMLLASFATSLTHLYLSIGLSSGWA L 120
Db 61 AVQQFGSPVGSALSTKFGPRPVVMTGGILALGMLLASFATSLTHLYLSIGLSSGWA L 120
QY 121 TFAPTLACLSCYFSRRSLATGLALTGVGLSSFTFA PFQWMLSHYAMRGSLLVSA LSL 180
Db 121 TFAPSLACLSCYFSRRSLATGLALTGVGLSSFTFA PFQWMLSHYAMRGSLLVSA LSL 180
QY 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHNGPFLRYTVALTLINTGYFIPLYLHV AH 240
Db 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHNGPFLRYTVALTLINTGYFIPLYLHV AH 240
QY 241 LQDLWDPLPAFAFLISVA ISDLVGRVSGWLGDAVPGPVTRLMLMTTLTGVS LALFPV 300
Db 241 LQDLWDPLPAFAFLISVA ISDLVGRVSGWLGDAVPGPVTRLMLMTTLTGVS LALFPV 300

Db	241	LQDLWDPLPAFLLSVAISDLVGRVSGMLGDVPGPVTRLMLWTTLTGVSALFPV	300
QY	301	AOAPTALVALAVAYGFTSGALAPLAFSVLPGLIGTRRIYCGILQMIESIGGLGPPLS	360
Db	301	AQAPTALVALAVAYGFTSGALAPLAFSVLPGLIGTRRIYCGILQMIESIGGLGPPLS	360
QY	361	GYLRDVSGNYTASFVWAGAFLLSGGILLTLPHFFCFSTTSGPQDLVTEALDTKVPPLK	420
Db	361	GYLRDVSGNYTASFVWAGAFLLSGGILLTLPHFFCFSTTSGPQDLVTEALDTKVPPLK	420
QY	421	EGLE 424	
Db	421	EGLE 424	
RESULT 10			
ID	AD116802		
XX	AD116802	standard; protein; 427 AA.	
AC	AD116802;		
XX			
DT	15-APR-2004	(first entry)	
XX			
DE	Human NOVX	protein homologue SegID 338.	
XX			
KW	human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;		
KW	inflammation; autoimmune disorder; allergy; blood disorder;		
KW	acquired immunodeficiency syndrome; AIDS; obesity; asclma;		
KW	immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;		
KW	Alzheimer's disease; infection; str.		
OS			
XX	Homo sapiens.		
PN	WO200268649-A2;		
XX			
PD	06-SEP-2002.		
PF	31-JAN-2002;	2002WO-US002785.	
XX			
PR	31-JAN-2001;	2001US-0265395P.	
PR	31-JAN-2001;	2001US-0265412P.	
PR	31-JAN-2001;	2001US-0265514P.	
PR	31-JAN-2001;	2001US-0265517P.	
PR	02-FEB-2001;	2001US-0266406P.	
PR	05-FEB-2001;	2001US-0266767P.	
PR	07-FEB-2001;	2001US-0266975P.	
PR	08-FEB-2001;	2001US-0267057P.	
PR	09-FEB-2001;	2001US-0267459P.	
PR	15-FEB-2001;	2001US-0267823P.	
PR	26-FEB-2001;	2001US-0268974P.	
PR	27-FEB-2001;	2001US-0271664P.	
PR	02-FEB-2001;	2001US-0271839P.	
PR	02-MAR-2001;	2001US-0271855P.	
PR	02-MAR-2001;	2001US-0272788P.	
PR	14-MAR-2001;	2001US-0273046P.	
PR	14-MAR-2001;	2001US-0275925P.	
PR	14-MAR-2001;	2001US-0275947P.	
PR	15-MAR-2001;	2001US-0275950P.	
PR	15-MAR-2001;	2001US-0275989P.	
PR	16-MAR-2001;	2001US-0276448P.	
PR	16-MAR-2001;	2001US-0276450P.	
PR	16-MAR-2001;	2001US-0276397P.	
PR	20-MAR-2001;	2001US-0276768P.	
PR	26-MAR-2001;	2001US-0278652P.	
PR	26-MAR-2001;	2001US-0278775P.	
PR	29-MAR-2001;	2001US-0278778P.	
PR	29-MAR-2001;	2001US-0279882P.	
PR	30-MAR-2001;	2001US-0279884P.	
PR	11-APR-2001;	2001US-0280147P.	
PR	11-APR-2001;	2001US-0282992P.	
PR	20-APR-2001;	2001US-0283083P.	
PR	20-APR-2001;	2001US-0285133P.	
PR	23-APR-2001;	2001US-0285749P.	

PR 03-MAY-2001; 2001US-0288327P.
 PR 03-MAY-2001; 2001US-0288504P.
 PR 29-MAY-2001; 2001US-0294047P.
 PR 30-MAY-2001; 2001US-0294473P.
 PR 08-JUN-2001; 2001US-0296964P.
 PR 18-JUN-2001; 2001US-0298959P.
 PR 19-JUN-2001; 2001US-0299324P.
 PR 13-AUG-2001; 2001US-0312020P.
 PR 16-AUG-2001; 2001US-0312889P.
 PR 16-AUG-2001; 2001US-0312908P.
 PR 21-AUG-2001; 2001US-0313390P.
 PR 28-AUG-2001; 2001US-0315470P.
 PR 31-AUG-2001; 2001US-0316447P.
 PR 07-SEP-2001; 2001US-0318115P.
 PR 07-SEP-2001; 2001US-0318118P.
 PR 12-SEP-2001; 2001US-0318740P.
 PR 19-SEP-2001; 2001US-0323379P.
 PR 18-OCT-2001; 2001US-0330245P.
 PR 18-OCT-2001; 2001US-0330308P.
 PR 14-NOV-2001; 2001US-03332701P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Tcherner VT, Spytek KA, Zerhusen BD, Patnirajan M, Shimkets RA;
 PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
 PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
 PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
 XX
 DR WPI; 2002-706998/76.
 XX
 PT New NOXV polypeptides and nucleic acids, useful for preventing or
 PT treating NOXV-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Disclosure; SEQ ID NO 338; 1498pp; English.
 XX

This invention relates to a novel nucleic acids, and encoded polypeptides thereof, which have properties related to the stimulation of biochemical or physiological responses in a cell, tissue, organ or organism. Specifically, it refers to the use of biologically active fragments for diagnostic and prognostic assays and furthermore in the treatment of diverse pathological conditions. The present invention describes novel human and murine NOVX proteins, as well as methods to modulate their expression using antisense oligos, ribozymes and peptide nucleic acids. The NOVX polypeptides, polynucleotides and antibodies are useful in treating or preventing NOVX-associated disorders, e.g. cardiomyopathy, atherosclerosis, cancer and diabetes. Furthermore, they may be used in treating or preventing diseases such as inflammation, autoimmune disorders, allergies, blood disorders, acquired immunodeficiency syndrome (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy and epilepsy. Accordingly, these molecules have many activities including cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic, haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic, antiasthmatic, nephrotropic, antiarthritic, hepatotropic, neuroprotective, nootropic, antibacterial, virucide, antiparasitic, relaxant and anticonvulsant. In addition, they are useful in screening assays to identify small molecules that modulate or inhibit, for example, neurogenesis, wound healing and angiogenesis. The nucleic acids are also used as in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. This polypeptide is a homologue of a human NOVX protein of the invention.

```

SQ      Sequence 427 AA;

Query Match
Best Local Similarity      91.9%; Score 2142.5; DB 5; Length 427;
Matches 422; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      1 MARRTEPPDGGWG-RVVVLSAFFQSALVEGYLRSFGVFFVEFVAAFEQAAARVSWIASIG 59
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1 MARRTEPPDGGWGXVVVLSAFFQSAIVFGVLRSGVFVFVEFVAAFEQAAARVSWIASIG 60

```


CC assays to identify small molecules that modulate or inhibit, for example, CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also used as in chromosome mapping, tissue typing, preventive medicine and CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein of the invention.

XX Sequence 427 AA;

Query Match 91.9%; Score 2142.5; DB 5; Length 427;
Best Local Similarity 99.3%; Pred. No. 3e-188;
Matches 422; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MARRTPEPPDGGWG-RVVVLSAFQSAIVFGVLRSGVFVEFEVAAFEEOARVSWIASIG 59
Db 1 MARRTPEPPDGGWGXVVVLSAFQSAIVFGVLRSGVFVEFEVAAFEEOARVSWIASIG 60
QY 60 IAVQOFGSPVGSALSTKFGPRPVMTGILAAIGMLASFATSLTHLYLSIGLSSGWA 119
Db 61 IAVQOFGSPVGSALSTKFGPRPVMTGILAAIGMLASFATSLTHLYLSIGLSSGWA 120
QY 120 LTFAPTLACLSYFSRRSLATGLATGVGLSSFTFAPFEQWMLSHYAWRGSLILVSALS 179
Db 121 LTFAPTLACLSYFSRRSLATGLATGVGLSSFTFAPFEQWMLSHYAWRGSLILVSALS 180
QY 180 LHLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFLRYTVALTLINTGYFIPLYHLVA 239
Db 181 LHLVACGALLRPPSLAEDPAVGGPRAQLTSLHHGPFLRYTVALTLINTGYFIPLYHLVA 240
QY 240 HLODLWDPLPAFLLSVAISDLVGRVSGWLGDAVGPVTRLMLMTLTGVSIALFP 299
Db 241 HLODLWDPLPAFLLSVAISDLVGRVSGWLGDAVGPVTRLMLMTLTGVSIALFP 300
QY 300 VAQAPTLVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGLGLOMIESIGLLGPPL 359
Db 301 VAQAPTLVALAVAYGFTSGALAPLAFSVLPBELIGTRRIYCGLGLOMIESIGLLGPPL 360
QY 360 SGYLRDVSGNYTASFVVAGAFLLSGGILLTLPHFPCFSTTSGPQDLVTEALDTKVPPLP 419
Db 361 SGYLRDVSGNYTASFVVAGAFLLSGGILLTLPHFPCFSTTSGPQDLVTEALDTKVPPLP 420
2Y 420 KEGLE 424
Db 421 KEGLE 425

RESULT 12

AD116766
ID AD116766 standard; protein; 426 AA.

AC AD116766;

CX 15-APR-2004 (first entry)

Human NOVX protein to treat human pathological conditions SeqID302.

human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
inflammation; autoimmune disorder; allergy; blood disorder;
acquired immunodeficiency syndrome; AIDS; obesity; asthma;
immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;
cytostatic; cardiant; antiinflammatory; immunosuppressive; antiallergic;
haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; anorectic;
antiaschmatic; nephrotropic; antiarthritic; hepatotropic;
neuroprotective; nootropic; antibacterial; virucide; antiparasitic;
relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis;
chromosome mapping; tissue typing; pharmacogenomic.

S Homo sapiens.

N WO200268649-A2.

D 06-SEP-2002.

PF 31-JAN-2002; 2002WO-US002785.

PR 31-JAN-2001; 2001US-0265395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266767P.
PR 07-FEB-2001; 2001US-0266975P.
PR 08-FEB-2001; 2001US-0267057P.
PR 09-FEB-2001; 2001US-0267459P.
PR 15-FEB-2001; 2001US-0267823P.
PR 26-FEB-2001; 2001US-0268974P.
PR 27-FEB-2001; 2001US-0271664P.
PR 02-MAR-2001; 2001US-0271839P.
PR 02-MAR-2001; 2001US-0272788P.
PR 14-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 15-MAR-2001; 2001US-0275989P.
PR 16-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 20-MAR-2001; 2001US-0276768P.
PR 26-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 29-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 30-MAR-2001; 2001US-0279884P.
PR 11-APR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 20-APR-2001; 2001US-0283083P.
PR 23-APR-2001; 2001US-0285133P.
PR 03-MAY-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 29-MAY-2001; 2001US-0288504P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 14-NOV-2001; 2001US-0332701P.

(CURA-) CURAGEN CORP.

Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shinkets RA,
Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE,
Gerlach VL, Taupier RJ, Gusev VY, Coleman SD, Wolenc AR, Pena CEA,
Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE,
WPI; 2002-706998/76.
N-PSDB; AD116765.

New NOVX polypeptides and nucleic acids, useful for preventing or
treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
pharmacogenomics.

Claim 1; SEQ ID NO 302; 1498bp; English.

This invention relates to a novel nucleic acids, and encoded polypeptides thereof, which have properties related to the stimulation of biochemical or physiological responses in a cell, tissue, organ or organism. Specifically, it refers to the use of biologically active fragments for diagnostic and prognostic assays and furthermore in the treatment of diverse pathological conditions. The present invention describes novel human and murine NOVX proteins, as well as methods to modulate their expression using antisense oligos, ribozymes and peptide nucleic acids. The NOVX polypeptides, polynucleotides and antibodies are useful in treating or preventing NOVX-associated disorders, e.g. cardiomyopathy, atherosclerosis, cancer and diabetes. Furthermore, they may be used in treating or preventing diseases such as inflammation, autoimmune disorders, allergies, blood disorders, acquired immunodeficiency syndrome (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy and epilepsy. Accordingly, these molecules have many activities including cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic, haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic, antiasthmatic, nephrotropic, antiarthritic, hepatotropic, neuroprotective, nootropic, antibacterial, virucide, antiparasitic, relaxant and anticonvulsant. In addition, they are useful in screening assays to identify small molecules that modulate or inhibit, for example, neurogenesis, wound healing and angiogenesis. The nucleic acids are also used as in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. This polypeptide is a human NOVX protein of the invention.

Sequence 426 AA;

Query Match	91.9%;	Score 2142;	DB 5;	Length 426;
Best Local Similarity	99.1%;	Pred. No. 3.3e-188;		
Matches 420; Conservative	2;	Mismatches 2;	Indels 0;	Gaps 0;

2y 1 MARRTEPDGCGWRVVVLSAFEQSALVEGVRSFGVFVEBFVNAFEEQARVSWIASIGI 60
|||||
1 MARRTEPDGCGWVVVLSAFEQSALVEGVRSFGVFVEBFVNAFEEQARVSWIASIGI 60

2y 61 AVQQTGSPVGSALSTKFGPRPVMTGILALGMLLASFATSLTHLYLSIGLLSGSGNAL 120

2b 61 AVQQTGSPVGSALSTKFGPRPVMTGILALGMLLASFATSLTHLYLSIGLLSGSGNAL 120

2y 121 TEAPTLACLSCTFSRRSLATGALTGVLSSTFPAPFQWLISHYAMRGSLLVSAISL 180
|||||
2b 121 TEAPTLACLSCTFSRRSLATGALTGVLSSTFPAPFQWLISHYAMRGSLLVSAISL 180
|||||

2y	181 HLVAAGALLRPPSLAEDPAVGCPRAQDLSLHGGPFLRYTVALTLINTGYFIPLYHLVAH 240
Db	181 HLVAAGALLRPPSLAEDPAVGCPRAQDLSLHGGPFLRYTVALTLINTGYFIPLYHLVAH 240

2y 241 LODLDMDPLPAAFLLSVVAISDLGVRVSGWLGDAPPGPVTRLMLMTTLTGVSALAEFV 300
|||||
Db 241 LODLDMDPLPAAFLLSVVAISDLGVRVSGWLGDAPPGPVTRLMLMTTLTGVSALAEFV 300

QY 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGLLQWIESIGLLGPPLS 360
|||||
db 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGLLQWIESIGLLGPPLS 360
|||||

```

QY      361  GYLRLDVSGNYLTASFVVVAGAFLLSGSGILLTPHFFCFSTTTTSGQDVLTEALDTKVPLPK 4200
        |||||:|||||
DB      361  GYLRLDVIGNYLTASFVVVAGAFLLSGSGILLTPHFFCSTTTTSGQDVLTEALDTKVPLPK 4200

```

QY	421	EGLF	424
Db	421	EGLF	424

RESULT 13	
AD116768	
ID	AD116768 standard; protein; 426 AA.
XX	
AC	
XX	AD116768;

DT 15-APR-2004 (first entry)

Human NOVX protein to treat human pathological conditions SeqID304.

human; NOX; cardiomyopathy; atherosclerosis; cancer; diabetes; inflammation; autoimmune disorder; allergy; blood disorder; acquired immunodeficiency syndrome; AIDS; obesity; asthma; immunoglobulin (Ig) A nephropathy; cirrhosis; arthritis; Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy; Alzheimers disease; infection; stroke; muscular dystrophy; epilepsy; cytostatic; cardiant; antiinflammatory; immunosuppressive; antiallergic; haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; anorectic; antiasthmatic; nephrotropic; antiarthritic; hepatotropic; neuroprotective; nootropic; antibacterial; virucide; antiparasitic; relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.

Os Homo sapiens.

PN WO200268649-A2.

PD 06-SEP-2002.

PF 31-JAN-2002; 2002WO-US002785.

PR 31-JAN-2001; 2001US-0265395P.

PR 31-JAN-2001; 2001US-0265514P.

PR 02-FEB-2001; 2001US-0266406P.

PR 07-FEB-2001; 2001US-0266975P.

PR 08-FEB-2001; 2001US-0267459P.

PR 15-FEB-2001; 2001US-0268974P.

PR 27-FEB-2001; 2001US-0271839P.

PR 02-MAR-2001; 2001US-0272788P.

PR 14-MAR-2001; 2001US-0275925P.

PR 14-MAR-2001; 2001US-0275950P.

PR 15-MAR-2001; 2001US-0276448P.

PR 16-MAR-2001; 2001US-0276397P.

PR 20-MAR-2001; 2001US-0278652P

PR 26-MAR-2001; 2001US-0278778P.
20 MAR 2001 2001US-0278882P

PR 29-MAR-2001; 2001US-0279888AP.
 20 MAR 2001 2001US-0380147P

PR 11-APR-2001; 2001US-0282992P
 DP 11-APR-2001; 2001US-0283083P

PR 20-APR-2001; 2001US-0285133P
 PR 23-APR-2001; 2001US-0285749P

PR 03-MAY-2001; 2001TTS-0288504P
PR 03-MAY-2001; 2001TTS-0288504P

PR 30-MAY-2001: 2001US-0294473P
PK 29-MAY-2001: 2001US-0294473P

18-JUN-2001: 2001US-0298959P

PR 13-AUG-2001; 2001US-0312020P.

PR 16-AUG-2001; 2001US-0312908P

PR 28-AUG-2001; 2001US-0315470P

PR 07-SEP-2001; 2001US-0318115P

PR 07-SEP-2001; 2001US-0318118P

30-MAY-2001; 2001US-0294473P.
08-JUN-2001; 2001US-0296964P.
18-JUN-2001; 2001US-0298959P.
19-JUN-2001; 2001US-0299324P.
13-AUG-2001; 2001US-0312020P.
16-AUG-2001; 2001US-0312889P.
16-AUG-2001; 2001US-0312908P.
21-AUG-2001; 2001US-0313930P.
28-AUG-2001; 2001US-0315470P.
31-AUG-2001; 2001US-0316447P.
07-SEP-2001; 2001US-0318115P.
07-SEP-2001; 2001US-0318118P.
12-SEP-2001; 2001US-0318740P.
19-SEP-2001; 2001US-0323379P.
18-OCT-2001; 2001US-0330245P.
18-OCT-2001; 2001US-0330308P.
14-NOV-2001; 2001US-0332701P.

(TCHE/) TCHERNEV V T.
(SPYT/) SPYTEK K A.
(ZERH/) ZERHUSEN B D.
(PAT/) PATTURAJAN M.
(SHIM/) SHIMKETS R A.
(LILL/) LI J.
(GANG/) GANGOLLI E A.
(PADI/) PADIGARU M.
(ANDE/) ANDERSON D W.
(RAST/) RASTELLI L.
(MILL/) MILLER C E.
(GERL/) GERLACH V.
(TAUP/) TAUPIER R J.
(GUSE/) GUSEV V Y.
(COLM/) COLMAN S D.
(WOLE/) WOLENC A R.
(PENA/) PENNA C E A.
(FURT/) FURTAKE K.
(GROS/) GROSSE W M.
(ALSO/) ALSCBROOK J P.
(LEPL/) LEPLEY D M.
(RIEG/) RIEGER D K.
(BURG/) BURGESS C E.

Tchernev VT, Spytek KA, Zehrusen BD, Patturajan M, Shimkets RA; Li J, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE; Gerlach V, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA; Furtae K, Grosse WM, Alscbrook JP, Lepley DM, Rieger DK, Burgess CE; WPI; 2004-180039/17.
N-PSDB; ADN42419.

Isolated NOVX polypeptides and polynucleotides, useful for preventing diagnosing and/or treating cancer, diabetes, cardiomyopathy and atherosclerosis.

Claim 1; SEQ ID NO 302; 1309pp; English.

The invention relates isolated 162 NOVX polypeptides (NOV1-NOV99, including splice variants) and the nucleic acids (NA) that encode them. Also included are the mature NOVX proteins (and their encoding polynucleotides), a vector comprising NOVX NA, a cell comprising the vector, an antibody that binds immunospecifically to NOVX, determining the presence or amount of NOVX in a sample, determining the presence or amount of NOVX NA in a sample, identifying an agent that binds to NOVX, modulating the activity of NOVX, treating or preventing a NOVX-associated disorder, determining the presence of or predisposition to a disease associated with altered levels of NOVX and treating a pathological state in a mammal comprising administering a polypeptide which is at least 95% identical to NOVX (or fragment). NOVX and NA may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate expression and activity of NOVX (e.g. cancer, diabetes, cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and antagonists may also be used to down regulate expression and activity of NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for

detecting the presence of NOVX in samples (e.g. by enzyme linked immunosorbant assay (ELISA). The agents and methods may be used in this way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy and/or atherosclerosis. The present sequence represents a NOVX protein.

Sequence 426 AA;
Query Match 91.9%; Score 2142; DB 8; Length 426;
Best Local Similarity 99.1%; Pred. No. 3.3e-188;
Matches 420; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 MARRTEPPDGGWGRVVVLSAFPQSALVFGVLRSGVFVEFVAAFEQARVSWIASIGI 60
1 MARRTEPPDGGWGVVVLSAFFQSALVFGVLRSGVFVEFVAAFEQARVSWIASIGI 60
61 AVQFGSPVGSALSTKEGPRPVMTGGILALGMLLASFATSLTHLYLSIGLSSGAWL 120
61 AVQFGSPVGSALSTKEGPRPVMTGGILALGMLLASFATSLTHLYLSIGLSSGAWL 120
121 TFAPTLACSCYFSRRSLATGLATGVLSSFTFAPFQWLLSHYAWRGSLIVSALS 180
121 TFAPTLACSCYFSRRSLATGLATGVLSSFTFAPFQWLLSHYAWRGSLIVSALS 180
181 HLVACGALLRPPSLAEDPAVCGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYHLVAH 240
181 HLVACGALLRPPSLAEDPAVCGPRAQLTSLHHGPFRLRYTVALTLINTGYFIPYHLVAH 240
241 LQDLMDPLPAFLLSVAISDLVGRVSGWIGDAVGPVTRLLMWTTLTGVSALFPV 300
241 LQDLMDPLPAFLLSVAISDLVGRVSGWIGDAVGPVTRLLMWTTLTGVSALFPV 300
301 AQAPTALVALAVAYFTSGALAPLAFSVLPPELIGTRITVCGLLQMIIESIGLGPPLS 360
301 AQAPTALVALAVAYFTSGALAPLAFSVLPPELIGTRITVCGLLQMIIESIGLGPPLS 360
361 GYLKDVSGNYTASFVAVAGAFILSGSGLLTLPHFFCFSTTTSGPQDLVTEALDTKVPLPK 420
361 GYLKDVSGNYTASFVAVAGAFILSGSGLLTLPHFFCFSTTTSGPQDLVTEALDTKVPLPK 420
421 EGLE 424
421 EGLE 424
421 EGLE 424

RESULT 15
ADN42422
ID ADN42422 standard; protein; 426 AA.

ADN42422;
ADN42422;
17-JUN-2004 (first entry)
Human novel proteinNOV 96b.
Human; NOVX; cancer; diabetes; cardiomyopathy; atherosclerosis.
Homo sapiens.
US2004033493-A1.
19-FEB-2004.
31-JAN-2002; 2002US-00072012.
31-JAN-2001; 2001US-0265395P.
31-JAN-2001; 2001US-0265412P.
31-JAN-2001; 2001US-0265514P.
31-JAN-2001; 2001US-0265517P.
02-FEB-2001; 2001US-0266406P.
05-FEB-2001; 2001US-0266767P.
07-FEB-2001; 2001US-0266975P.
08-FEB-2001; 2001US-0267057P.
08-FEB-2001; 2001US-0267459P.

PR 09-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 02-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 14-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313930P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.

XX (TCHE/) TCHERNEV V T.
PA (SPYT/) SPYTEK K A.
PA (ZERH/) ZERHUSEN B D.
PA (PAT/) PATURAJAN M.
PA (SHIM/) SHIMKETS R A.
PA (LILL/) LI L.
PA (GANG/) GANGOLLI E A.
PA (PADI/) PADIGARU M.
PA (ANDE/) ANDERSON D W.
PA (RAST/) RASTELLI L.
PA (MILL/) MILLER C E.
PA (GERL/) GERLACH V.
PA (TAUP/) TAUPIER R J.
PA (GUSE/) GUSEV V Y.
A (COLM/) COLMAN S D.
A (WOLE/) WOLENC A R.
A (PENA/) PENNA C E A.
A (FURT/) FURTAK K.
A (GROS/) GROSSE W M.
A (ALSO/) ALSOBROOK J P.
A (LEPL/) LEPLEY D M.
A (RIEG/) RIEGER D K.
A (BURG/) BURGESS C E.

I Tchernev VT, Spytek KA, Zerhusen BD, Paturajan M, Shimkets RA;
I Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
I Gerlach V, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;

PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
DR MPI; 2004-180039/17.
DR N-PSDB; ADN42421.
XX
PT Isolated NOVX polypeptides and polynucleotides, useful for preventing
PT diagnosing and/or treating cancer, diabetes, cardiomyopathy and
PS atherosclerosis.

Claim 1; SEQ ID NO 304; 1309PP; English.

CC The invention relates isolated 162 NOVX polypeptides (NOV1-NOV99,
CC including splice variants) and the nucleic acids (NA) that encode them.
CC Also included are the mature NOVX proteins (and their encoding
CC polynucleotides), a vector comprising NOVX NA, a cell comprising the
CC vector, an antibody that binds immunospecifically to NOVX, determining
CC the presence or amount of NOVX in a sample, identifying an agent that binds to NOVX,
CC amount of NOVX NA in a sample, identifying an agent that binds to NOVX,
CC modulating the activity of NOVX, treating or preventing a NOVX-associated
CC disorder, determining the presence of or predisposition to a disease
CC associated with altered levels of NOVX and treating a pathological state
CC in a mammal comprising administering a polypeptide which is at least 95%
CC identical to NOVX (or fragment). NOVX and NA may be used in the
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression and activity of NOVX (e.g. cancer, diabetes,
CC cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and
CC antagonists may also be used to down regulate expression and activity of
CC NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for
CC detecting the presence of NOVX in samples (e.g. by enzyme linked
CC immunosorbant assay (ELISA). The agents and methods may be used in this
CC way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy
CC and/or atherosclerosis. The present sequence represents a NOVX protein.

XX Sequence 426 AA;

Query Match 91.9%; Score 2142; DB 8; Length 426;
Best Local Similarity 99.1%; Pred. No. 3.3e-188;
Matches 420; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MARRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOARVSWIASIGI 60
Db 1 MARRTEPPDGGWGWVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOARVSWIASIGI 60
QY 61 AVQQFGSPVGSALSTKFGPRPVVMTGGIHALGMLIASFATSLTHLYSIGLSSGSWAL 120
Db 61 AVQQFGSPVGSALSTKFGPRPVVMTGGIHALGMLIASFATSLTHLYSIGLSSGSWAL 120
QY 121 TFAPTLACISCFYSRRSLATGLATGVGLSSFTFAPFQWLSHYAWRGSLLVSALSL 180
Db 121 TFAPTLACISCFYSRRSLATGLATGVGLSSFTFAPFQWLSHYAWRGSLLVSALSL 180
QY 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHNGPFLRYTVALTLINTGYFIPYLHVAH 240
Db 181 HLVACGALLRPPSLAEDPAVGGPRAQLTSLHNGPFLRYTVALTLINTGYFIPYLHVAH 240
QY 241 LQDLWDPLPAFLISVAISDLVGRVSVGWLGDVAVGPVTRLLMLMTTLTGVSALFPV 300
Db 241 LQDLWDPLPAFLISVAISDLVGRVSVGWLGDVAVGPVTRLLMLMTTLTGVSALFPV 300
QY 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPGLIGTRRIYCGLLQMIIESIGLGPPLS 360
Db 301 AQAPTALVALAVAYGFTSGALAPLAFSVLPGLIGTRRIYCGLLQMIIESIGLGPPLS 360
QY 361 GYLKDVSGNYTASFVAGAFLLSSGILTLPHFCCSTTSGPQDLVTEALDTKVPLPK 420
Db 361 GYLKDVSGNYTASFVAGAFLLSSGILTLPHFCCSTTSGPQDLVTEALDTKVPLPK 420
QY 421 EGLE 424
Db 421 EGLE 424

Thu Sep 7 14:12:15 2006

job time : 199 secs

us-10-654-428-2.rag

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2006, 13:07:31 ; Search time 50 Seconds

(without alignments)
800.030 Million cell updates/sec

Title: US-10-654-428-2
Perfect score: 2331
Sequence: 1 MARRTEPPDGGWGRVVLSA.....SLTAPGLLPRLGLHRTTVP 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*

1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5_COMB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6_COMB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7_COMB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H_COMB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pep:*
7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	608	26.1	478	2	US-09-949-016-6543 Sequence 6543, Ap
2	605.5	26.0	485	2	US-09-949-016-10465 Sequence 10465, A
3	599.5	25.7	523	2	US-09-949-016-6537 Sequence 6537, Ap
4	593	25.4	527	2	US-09-949-016-9381 Sequence 9381, Ap
5	591.5	25.4	473	2	US-09-949-016-11735 Sequence 11735, A
6	540.5	23.2	500	2	US-09-976-594-419 Sequence 419, App
7	490	21.0	505	2	US-09-949-016-6538 Sequence 6538, Ap
8	490	21.0	521	2	US-09-949-016-8809 Sequence 8809, Ap
9	473.5	20.3	487	2	US-10-074-547-5 Sequence 5, Appli
10	473.5	20.3	492	2	US-09-949-016-10677 Sequence 10677, A
11	468	20.1	510	2	US-10-074-547-2 Sequence 2, Appli
12	417.5	17.9	634	2	US-09-949-016-7542 Sequence 7542, Ap
13	398.5	17.1	611	2	US-10-074-547-4 Sequence 4, Appli
14	376	16.1	480	2	US-09-270-767-44472 Sequence 44472, A
15	350	15.0	429	2	US-09-328-352-4392 Sequence 4392, Ap
16	345	14.8	394	2	US-10-154-419-28 Sequence 28, Appli
17	291	12.5	403	2	US-09-252-991A-30953 Sequence 30953, A
18	246	10.6	151	2	US-09-270-767-32358 Sequence 32358, A
19	246	10.6	151	2	US-09-270-767-47575 Sequence 47575, A
20	240	10.3	443	2	US-09-248-796A-20821 Sequence 20821, A
21	236	10.1	407	2	US-09-489-039A-10855 Sequence 10855, A
22	234	10.0	348	2	US-09-270-767-32307 Sequence 32307, A
23	234	10.0	348	2	US-09-270-767-47524 Sequence 47524, A
24	226	9.7	336	2	US-10-094-749-1996 Sequence 1996, Ap
25	224	9.6	534	2	US-09-252-991A-27300 Sequence 27300, A
26	213.5	9.2	228	2	US-09-973-278-160 Sequence 160, App

27	213.5	9.2	229	2	US-09-227-357-237	Sequence 237, App
28	191	8.2	467	2	US-09-489-039A-9693	Sequence 9693, Ap
29	185	7.9	482	2	US-09-489-039A-9528	Sequence 9528, Ap
30	183.5	7.9	463	2	US-09-489-039A-10827	Sequence 10827, A
31	182	7.8	450	2	US-09-252-991A-25019	Sequence 25019, A
32	176	7.6	417	2	US-09-489-039A-7357	Sequence 7357, Ap
33	174.5	7.5	447	2	US-09-248-796A-20825	Sequence 20825, A
34	173.5	7.4	500	2	US-09-252-991A-25053	Sequence 25053, A
35	171.5	7.4	555	2	US-09-252-991A-31521	Sequence 31521, A
36	171	7.3	345	2	US-09-248-796A-20831	Sequence 20831, A
37	170.5	7.3	397	2	US-09-489-039A-12662	Sequence 12662, A
38	170	7.3	381	2	US-09-248-796A-20832	Sequence 20832, A
39	170	7.3	421	2	US-09-489-039A-7699	Sequence 7699, Ap
40	170	7.3	495	2	US-09-328-352-5768	Sequence 19278, A
41	169.5	7.3	408	2	US-09-543-681A-6289	Sequence 5768, Ap
42	169.5	7.3	420	2	US-09-489-039A-9508	Sequence 6289, Ap
43	169.5	7.3	462	2	US-09-489-039A-13633	Sequence 9508, Ap
44	169	7.3	434	2	US-09-489-039A-13633	Sequence 13633, A
45	168	7.2	385	2	US-09-270-767-42231	Sequence 42231, A

ALIGNMENTS

RESULT 1
US-09-949-016-6543
; Sequence 6543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6543
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-6543

Query Match	26.1%; Score 608; DB 2; Length 478;
Best local Similarity	32.8%; Pred. No. 2.3e-50;
Matches	152; Conservative 79; Mismatches 180; Indels 52; Gaps 7;
QY	7 PPDGGRVVVLSAFFQSALVFGVLSFGVFFVEFVAAPFEQAAARVSWIASIGIAVQQFG 66
DB	13 PPDGGMGIWVGAAFTSIGFSYAFPKAVTVFKEIQIIFHTTYSEIAWISSIMLAVMYAG 72
QY	67 SPVGSALSTKFGPRPVVMTGIIAALGMLLASFATSLTHULYSIGLLSGSGWALTFAPTL 126
DB	73 GPVSSVLVNNKYGSRPVVIAGLCLGLMVLASFSSSVQLYLTMGFITGLGLAFNLQFAL 132
QY	127 ACISCFYSRRSLATGLALTGVLSFTFAPFFQWLISHYAWRGSLLVSAISLHLVACG 186
DB	133 TIIGKYFYRKRPMPANGLAAGSPVFLSSLAPFNQYLFNTFGWKGSFLILGSLILNACVAG 192
QY	187 ALLRP-----PSLAEDPAVGGPRAQLT-----SLHHGPFLRY 219
DB	193 SLMRPLGPNOTTSKSKNKTGKTEDDS--SPKKIKTKKSTWEKVNKYLDPSLFRGRGLIY 250
QY	220 TVALTLINTGYFIPYLHVAHLQDLWDPLPAAPFLSVVAISDLVGRVSWGVLGDA--VP 277
DB	251 LSGNVIMFLGFFAPLIFLAPYAKDQIDEYSAFLLSVMAFVDMFARPSVGLIANSKYIR 310

QY 278 GPVTRLMLWTTLTGVSLALFPVAQAPALVALAVAGFTSGALAPLAFSVLPeligTRR 337
Db 311 PRIQYFFSFAMFENGCHLLCPLAQDYTSLVAVFFGLGFSVSVLFETLMDLVGAPR 370
QY 338 IYCGGLLQMIESTIGLGLPPLSGYLRDVSNGNYTASFVAVAGAFLLSGGILLTLPHFCF 397
Db 371 FSSAVGLVTIVECGVLLGPPLAGKLVDTLGEYKMYMSCAIVVAVASVWL----- 421
QY 398 STTTSGPQDLVTEALDTKVPPLPKEGLEGL--NSTESGPESQS 438
Db 422 -----LIGNAINYRL-LAKERKEENAROKSRESEPLSKS 454

RESULT 2
US-09-949-016-10465
; Sequence 10465, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10465
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10465

Query Match 26.0%; Score 605.5; DB 2; Length 485;
Best Local Similarity 34.2%; Pred. No. 4.1e-50;
Matches 141; Conservative 73; Mismatches 167; Indels 31; Gaps 4;

QY 7 PPDGCGRNVVLSAFQSAIVFGVLSFGVFVEFVAAFEQARVSMIASIGIAVQOFG 66
Db 20 PPDGCGWIVVGAFAISIGFSYAFPKAVTVFFKEIQIFHTTYSEIAMISSIMLAVMYAG 79
QY 67 SPVGSALSTKFGPRPVVMTGGLIALGMLLASFATSLTHLYLSIGLLSGSGWALTFAPTL 126
Db 80 GPVSSVLVNKYSRPPVIVAGLLCCLGMVLASFSSSVQVLYLTMGFITGLAFNLQPAL 139
QY 127 AGLSCYFSRRSLATGLALTGVGLSSFTFAPFFQWLISHYAWRGSLLLVLSALSILHVAAG 186
Db 140 TIIGKYFYRKRPMANGLAMAGSPVFLSSLAPFNQYLFNTEGWMKGSFLILGSLLLNACVAG 199
QY 187 ALLRP-----PSLAEDPAVGGPRAQLT-----SLHNGPFLRY 219
Db 200 SLMKPLGPNQOTTSKSKNKTGKTEDDS--SPKKIKTKKSTWEKVNKYLDFSLFKRGLIY 257
QY 220 TVALTLINTGYFIPYHLVVAHLODLWDPLPAAPLISVAISDLVGRVSGWLGDA--VP 277
Db 258 LSGNVIMFLGFAPIIFLAPYAKDOGIDEYSAAFLLSVMAFVDMFARPSVGLIANSKYIR 317
QY 278 GPVTRLMLWTTLTGVSLALFPVAQAPALVALAVAGFTSGALAPLAFSVLPeligTRR 337
Db 318 PRIQYFFSFAMFENGCHLLCPLAQDYTSLVAVFFGLGFSVSVLFETLMDLVGAPR 377
QY 338 IYCGGLLQMIESTIGLGLPPLSGYLRDVSNGNYTASFVAVAGAFLLSGGILL 389
Db 378 FSSAVGLVTIVECGPVLGLPPLAGKLVDTLGEYKMYMSCAIVVAVASVWL 429

RESULT 3
US-09-949-016-6537

; Sequence 6537, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6537
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6537

Query Match 25.7%; Score 599.5; DB 2; Length 523;
Best Local Similarity 30.9%; Pred. No. 1.8e-49;
Matches 156; Conservative 83; Mismatches 181; Indels 85; Gaps 8;

QY 5 TEPDGGGRNVVLSAFQSAIVFGVLSFGVFVEFVAAFEQARVSMIASIGIAVQO 64
Db 16 TEPDGGGWAVAVSFFVEVFTYGIKTFGVFENDLMSFMSNSRISWISICVFVLT 75
QY 65 FGSVPGSALSTKFGPRPVVMTGGLIALGMLLASFATSLTHLYLSIGLLSGSGWALTFA 124
Db 76 FSAPLATVLSNRFGLRLVVMGLVSTGMVAASFQEVSHMYVAIGIISGLGYCFSLP 135
QY 125 TLACLSYFSRRSLATGLALTGVGLSSFTFAPFFQWLISHYAWRGSLLLVLSALSILHVA 184
Db 136 TVTILSQYFGKRSIVTAVASTGECAVFAFAPAIMALKERIGWRYSLLFVGLLQINVI 195
QY 185 CGALLRP-----PSL-----AEDPAVGG----- 202
Db 196 FGALLRPILIRGPASPPIVQENRKEAQYMLENEXTRTSIDSGVELLTSKPVPTHT 255
QY 203 -----PRAQLT-----SLHNGPFLRYTVALTLINTGYFIPYHL 237
Db 256 NLELEPKADMQOVLVKTSPRPSEKAPLLDFSILKEKSFICYALFGLFATLGFAPSLYI 315
QY 238 VAHLQDLMDPLPAAPLISVAISDLVGRVSGWLGDAVPQVTRLMLWTTLTGVSLAL 297
Db 316 IPLGISLGIDQDRAAFLLSTMAIAEYFGRIGAGFVLNREPDIRKIYIELICVILLTVSLFA 375
QY 298 FPAQAPALVALAVAGFTSGALAPLAFSVLPeli--GTRRIYCGLLQMIESTIGL 355
Db 376 FTFATEFWMGIMSCSIFFGFMVGTIGLTFHCLLKMMSWALQKSSAAGVIYIFIQSIAGLA 435
QY 356 GPPLSGYLRDVSNGNYTASFV-VAGAFLLSGGILLTLP-----HFFCFSTTSGPQD 406
Db 436 GPPLAGLVLDQSKIYSRAFYSCAGMALAAVCLALVRPCKMGLQGRHH-----SGETK 488
QY 407 LVTEALPTKVPPLPKEGLEGINSTE 431
Db 489 VVSHRGKTLQDIPEDFLEMDLAKNE 513

RESULT 4
US-09-949-016-9381
; Sequence 9381, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307

```

; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9381
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9381
```

Query Match 25.4%; Score 593; DB 2; Length 527;
Best Local Similarity 31.1%; Pred. No. 7.6e-49;
Matches 157; Conservative 83; Mismatches 181; Indels 84; Gaps 10;

```

QY 5 TEPPDGGWGRVVLSAFFQSALVFGVLSFGVFVEFVAAFEEOAARVSWIASIGIAVQQ 64
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 19 TEVPDGGWGAVAVSFFFEVFTYGIKTGFGVFENDLMSFNESNSRISWISICVFLVLT 78
QY 65 EGSFVGSAALSTKFGPRPVMTGILALGLMLASFATSLTHLYLSIGLLSGSGWALTFAF 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 79 ESAPLATVLSNRFGHRVLVVLGGLVSTGVAASFQEVSHMYVAIGIISGLGCFSEFLP 138
QY 125 TLACLSCYFSRRSLATGLATGVGLSSFTFAPFQWLSSHVAMRGSLLSALSLHLVA 184
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 139 TVTTLSEQYFGKRISIVTAVASTGCEFAVAFAPAIMALKERIGWRYSLTFVGLQINIVI 198
QY 185 CGALLRP-----PSL-----AEDPAVG----- 202
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 199 FGALLRPFIIRGPASPKIVIQENRKEAQYMLENEKRTSIDSIDSGVELTSPKVPPTH 258
QY 203 -----PRAQLTSL-----HHGFLRYTV-----ALTINT-GYFIPLYH 236
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 259 NLELEPKADMQVLVKTSPRPSEKAPLDFSIFEREKPYLVMLFGLFATIGFFAPSLY 318
QY 237 LVAIILQDLWDPLPAAFLLSVAISDLVGRVSGWLGDAVPGVTRLLMLMTLLTGVSIA 296
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 319 IIPGLISLIGIDQRAAFLLSTMAIAEVEFRIGAGFVLNREPIRKIYIELICVILLTVSLF 378
QY 297 LFPVAQAPFALVALAVAYGFTSGALAPLAFSVLPE--LIGTRRIYCGGLQMIESIGL 354
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 379 AFTFATEEWGLMSCSIFEGFMVGTIGTHIPLAEDDVVGIEKMSAAGVYIFIQSIAGL 438
QY 355 LGPPLSGYLRDVSGNYTASFV-VAGAFLLSGSGILLTL-----HFFCFSTTSGPOD 406
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 433 AGTPLAGLVDQSKIYSRAFYSCAAGMALAACLALVRPCKMGLCQHNN-----SGETK 492
QY 407 LVTEALDTKVPLPKEGLEGLNSTE 431
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 493 VVSHRGKTLQDIPEDFLEMDLAKNE 517
```

RESULT 5
US-09-949-016-11735
Sequence 11735, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11735
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11735
```

Query Match 25.4%; Score 591.5; DB 2; Length 473;
Best Local Similarity 33.7%; Pred. No. 9.1e-49;
Matches 141; Conservative 76; Mismatches 193; Indels 9; Gaps 4;

```

QY 8 PDGGWGRVVLSAFFQSALVFGVLSFGVFVEFVAAFEEOAARVSWIASIGIAVQQFGS 67
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24 PDGGWGAVALFGCFVITGFSYAFPKAVSVFFKELIQEFGIGYSDTAWISSILLMLXGTG 83
QY 68 PVGSALSTKFGPRPVMTGILALGLMLASFATSLTHLYLSIGLLSGSGWALTFAPTLA 127
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 84 PLCSVCVNRFGCRPVMLVGLFASLGVVAASFCSIIQVYLTGTGTLGLALNFQPSLI 143
QY 128 CLSCYFSRRSLATGLATGVGLSSFTFAPFQWLSSHVAMRGSLLSALSLHLVACGA 187
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 144 MLNRYFSKRRPMANGLAAGSPVFLCALSPDLQQLQDRYGMRGGLILGILLNCVCVAA 203
QY 188 LLRPPLAEDPAVGPRAL-----TSLHHCPLRYTVALTLINTGYFIPLYHLVAHLOD 243
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 204 LMRPLVTTAOPGSGPPRPSRRLDLISVFRDRGFVLVAVAAVMVLGFVPVFSYAKD 263
QY 244 LDWDPLPAAFLLSVAISDLVGRVSGW--LGDVAPGPVTRLLMLMTLLTGVSIALFPV 300
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 264 LGVPDTKAAFLTLTILGFIDIFARPAAGFVAGLGKVRPYSV-YHFSFMEFNGLADLAGST 322
QY 301 AQAPFALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGGLQMIESIGLLGPPLS 360
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 323 AGDYGLLVFCIFFGISYGMWGAQLQEVLMIAIVGTHKFSALGLVLMBAVAVLVGPPSG 382
QY 361 GYLRDVSGNYTASFVAVAGAFLLSGSGILLTLPHFFCFSTTSGPODVLTEALDTKVPLP 419
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 383 GKLDATHVYMYVFIAGAEVLTSSLILL-LGNFFCIRKKPKPEQPEVAAAEEKLRKP 440
```

RESULT 6
US-09-976-594-419
Sequence 419, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 419
LENGTH: 500
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6673549 2545475CD1
US-09-976-594-419

Query Match 23.2%; Score 540.5; DB 2; Length 500;
Best Local Similarity 29.2%; Pred. No. 8.9e-44;
Matches 147; Conservative 72; Mismatches 192; Indels 93; Gaps 11;

```

QY 7 PPDGSGWGRVVLSAFFQSALVFGVLSFGVFVEFVAAFEEOAARVSWIASIGIAVQQFG 66
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13 PPDGSGWGAVALVIGAFISIGFSYAFPEKSIITVFKEIEGIFHATTSEVSWISSIMLAVMYGG 72
```


[illegible]

RESULT 7
US-09-949-016-6538
; Sequence 6538, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6538
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6538

Query Match	21.0%;	Score 490;	DB 2;	Length 505;
Best Local Similarity	29.0%;	Pred. No. 7.3e-39;		
Matches 134;	Conservative 75;	Mismatches 213;	Indels 40;	Gaps 6;

[illegible]

```

Db      121 SFQSSITVLGFYFVRRRLVANALALASMGVSLGITLWPLLSRYLLENLUGWRGTFLVEGGIFL 180

QY      181 HLVACGALIR-----PESLAEDPAVGGPRAQLTSLHHGPF--LRYTVALTL 225
        |||::|      |||::|      |||::|      |||::|      |||::|
Db      181 HCCICGAIIRPVATSVAPETKECPBPPEPALGCLAACGRTIQRLAFDILRH----- 234

QY      226 INTGYFI-----PYLHLVAHLQDLWDPLPAAFILSVVAISDLVGRVSGW 271
        ||||:      |||::|      |||::|      |||::|      |||::|
Db      235 -NTGYCVIILGVMMWSVLGFPLPQVFLVFPYAMMHSVDEQQAALLISIIIGFSNIFLRPLAGL 293

QY      272 LG--DAVGEPVTRLMLWTTLTGVSLLALFPVAQAFTALVALAVAGFTSGALAPLAFSVL 329
        :|      |      |      |      |      |      |      |      |      |
Db      294 MAGRPAPASHRKYLFSLLALLNGLTNLNVCAASGDFWVLVGYCLAVSVMSGIGALIFQVL 353

QY      330 PELIGTRRIYCGLLQMIIESIGLLGPPLSGYLRDVSNGNYTASFVAVAGAFLLSGGILL 389
        :::      |||::|      :::      |||::|      :::      |||::|      :::      |||::|
Db      354 MDIVPMQDEPRALGLFTVLDSLAFILSPPLAGLLDATNNFSYVFYMWSSFFLISAA--LF 411

QY      390 TLPHFFCFSTTSGPQDLVTEALDTKYVLPKEGLEGLNSTE 431
        :|      |      |      |      |      |      |      |      |      |
Db      412 MGSFYLQKEQGQKQAVADALEKDLFLLEAKDGPQKQRSPE 453

```

```

RESULT 8
US-09-949-016-8809
; Sequence 8809, Application US/09949016
; Patent No. 6812339
;
GENERAL INFORMATION:
;
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241, 755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237, 768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231, 498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 8809
;
LENGTH: 521
;
TYPE: PRT
;
ORGANISM: Human
US-09-949-016-8809

```

Query Match	21.0%;	Score 490;	DB 2;	Length 521;
Best Local Similarity	29.0%;	Pred. No. 7.6e-39;		
Matches 134;	Conservative 75;	Mismatches 213;	Indels 40;	Gaps 6;

[illegible]

Db 310 MAGRPAPASHRKVYLFSLALLNLNLTNVCASGDFWLVGVCLAYSVSMGICALIFQVL 369
QY 330 PELIGTRRIYCGLLGLLOMIESIGLLGPPISGYLRDVSNGNYTASFVAGAFLLSGGILL 389
Db 370 MDIVPMQPPRALGLFTVLDGLAFLLISPLAGLLDATNMFSYFYMSSFFLISAA--LF 427
QY 390 TLPHPFCSTTSGPODLVTEALDTKVPLEKGEGLNSTE 431
Db 428 MGSFYALQKEQKQAVADALERDLFLEAKDGPGRORSPE 469

RESULT 9
US-10-074-547-5
; Sequence 5, Application US/10074547
; Patent No. 6989441
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Millennium Pharmaceuticals Inc.
; TITLE OF INVENTION: 25466, A Human Transporter and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MP12001-019P1RCPI(M)
; CURRENT APPLICATION NUMBER: US/10/074,547
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/269072
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 487
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-074-547-5

Query Match 20.3%; Score 473.5; DB 2; Length 487;
Best Local Similarity 26.6%; Pred. No. 2.8e-37;
Matches 122; Conservative 78; Mismatches 178; Indels 81; Gaps 5;
QY 5 TEPDGGWGRVVVLSAFFQSALVFGVRSFGVFVEFVAAFEEOAARVSIASIGIAVQ 64
Db 12 TKTLDDGGMWIVIHFFLVNVFVMGMTKPAIFVVFQEEFEGTSEQIGWIGSIMSLRF 71
QY 65 FGSPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLISGSWALTFFAP 124
Db 72 CAGPLVAICDILGEKTTSLGAFVVTGGYLISSWATSIPLCVTMGLLPGLGSAFLYQV 131
QY 125 TLACLSGFSRRSLATGLALTGVGLSSFTFAPFFQWLASHYAWRGSLLLVSALSLHLVA 184
Db 132 AAVVTTKYFKKRLALSTAIARSGMGL-TFLAPFTKFLIDLVDWTGALLFGAIALNLVP 190
QY 185 CGALLRPPSLAEDPAVG-----GPRQOLTSL----- 210
Db 191 SSMILRPIHIKSENNSGIKDKGSSLSAHGPEAHATETHCHETEESTIKDSTQKAGLPBK 250
QY 211 -----LHGP----- 225
Db 251 NLTVSONOSEEFYNGPNRNRLLLKSDEESDKVISWSCKQLFDISLFRNPFYIFTWGLL 310
QY 226 INTGYFIPIYHLVAHLQDLWDPLPAFLLSVAISDLVGRVSWGMLGDAVPGPVTRLLM 285
Db 311 SOLAYFIPTFHLVARAKTLGIDIMDASYLSVAGILETVSQIISGWADQNMWKYHYHK 370
QY 286 LMTTLTGVSLLAFVPAQAPTAVALAVAYGFTSGALAPLAFSVLPDELIGTRRIYCGILL 345
Db 371 SYLLLCGITNLAPLATTEPPLMTYTTICFAIFAGGYLALLLPVLVDLCRNSTVNRFLGLA 430
QY 346 OMIESIGLLGPPLSGYLRDVSNGNYTASFVAG-AFLLS 383
Db 431 SFPAGMAVLSPPIAGWLYDYQTYNGSFYFSGICYLLS 469

RESULT 10
US-09-949-016-10677

; Sequence 10677, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10677
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10677

Query Match 20.3%; Score 473.5; DB 2; Length 492;
Best Local Similarity 26.6%; Pred. No. 2.8e-37;
Matches 122; Conservative 78; Mismatches 178; Indels 81; Gaps 5;
QY 5 TEPDGGWGRVVVLSAFFQSALVFGVRSFGVFVEFVAAFEEOAARVSIASIGIAVQ 64
Db 17 TKTLDDGGMWIVIHFFLVNVFVMGMTKPAIFVVFQEEFEGTSEQIGWIGSIMSLRF 76
QY 65 FGSPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLISGSWALTFFAP 124
Db 77 CAGPLVAICDILGEKTTSLGAFVVTGGYLISSWATSIPLCVTMGLLPGLGSAFLYQV 136
QY 125 TLACLSGFSRRSLATGLALTGVGLSSFTFAPFFQWLASHYAWRGSLLLVSALSLHLVA 184
Db 137 AAVVTTKYFKKRLALSTAIARSGMGL-TFLAPFTKFLIDLVDWTGALLFGAIALNLVP 195
QY 185 CGALLRPPSLAEDPAVG-----GPRQOLTSL----- 210
Db 196 SSMILRPIHIKSENNSGIKDKGSSLSAHGPEAHATETHCHETEESTIKDSTQKAGLPBK 255
QY 211 -----LHGP----- 225
Db 256 NLTVSONOSEEFYNGPNRNRLLLKSDEESDKVISWSCKQLFDISLFRNPFYIFTWGLL 315
QY 226 INTGYFIPIYHLVAHLQDLWDPLPAFLLSVAISDLVGRVSWGMLGDAVPGPVTRLLM 285
Db 316 SOLAYFIPTFHLVARAKTLGIDIMDASYLSVAGILETVSQIISGWADQNMWKYHYHK 375
QY 286 LMTTLTGVSLLAFVPAQAPTAVALAVAYGFTSGALAPLAFSVLPDELIGTRRIYCGILL 345
Db 376 SYLLLCGITNLAPLATTEPPLMTYTTICFAIFAGGYLALLLPVLVDLCRNSTVNRFLGLA 435
QY 346 OMIESIGLLGPPLSGYLRDVSNGNYTASFVAG-AFLLS 383
Db 436 SFPAGMAVLSPPIAGWLYDYQTYNGSFYFSGICYLLS 474

RESULT 11
US-10-074-547-2
; Sequence 2, Application US/10074547
; Patent No. 6989441
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Millennium Pharmaceuticals Inc.
; TITLE OF INVENTION: 25466, A Human Transporter and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MP12001-019P1RCPI(M)
; CURRENT APPLICATION NUMBER: US/10/074,547
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/269072

;; PRIOR FILING DATE: 2001-02-15
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 510
;; TYPE: PRT
;; ORGANISM: homo sapiens
US-10-074-547-2

Query Match 20.1%; Score 468; DB 2; Length 510;
Best Local Similarity 25.2%; Pred. No. 1e-36;
Matches 119; Conservative 90; Mismatches 166; Indels 98; Gaps 10;

QY 9 DGGWARRVVVLSAFFQSALVEGVLSFGVFVEFVAAFEQAAARVSWIASIGIAVQFGSP 68
Db 29 DGGWAMMMVLSFFVHILIMSGMALGVNLVEMLEEFHQSRGLTAMVSSLSMGITLIVGP 88
QY -VGSALSTKFGPRPVMTGILALGMLLASFATSLTHLYLSIGLLSGSGWALTFAPTLA 127
Db 89 FIGLFINT-CGCRQTAIGGLVNSLGWLSAYAAVHYLFTTFGVAAGLSSGMALPAVV 147
QY 128 CISCYFSRRSLATGLALTGVGSSFTFAPFQWLLSHYAWRGSLLVLSALSLHVACGA 187
Db 148 MWGRYFQKRRLAQGLSTGTGEGFTLMTVLKYLCAEYGRNAMLIQAVSLNLCVCGA 207
QY 188 ILRPSLAEDPAVGGPRAQLTSLHHPFLRYT----- 220
Db 208 LMRPLSPGKNPNDPGEKDVRLPAHSTESVKSTGQGRTEKDGGLGNETLCDLQAGEC 267
QY 221 -----VALTLINT-----GYF----- 231
Db 268 PDQAGHRKNMICALRIKTVSWLTMVRKGFEDWYSGYFGTASLFTNRMFAFIWALFAY 327
QY 232 ----IPYLRH--VAHLODLDM--DPLPAFLLSVAISDLVGRVSGWLGDAVPGPVTRL 283
Db 328 SSFVIFPIHPEIVNLVNLSEQNDVFP--LTSIIAIVHIFGKVLGVIAIDLPCISVMNV 384
QY 284 LMLWTTLTGVSALFPVAQAPTALVALVAAYGFTSGALAPLAFSVLP---ELIGTRRIY 339
Db 385 FLANFTLVLSIFILPLMHTYAGLAVICALIGFSSG-----YFSLMPVTEDLVGIEHLA 439
QY 340 CGLJLQMIESIGLLGPPLSGYLKRDVSGNYTASFVAVAGAFLLSGGILLTLP 392
Db 440 NAYGIICANGISALLGPPFAGWIYDTQKYDFSFYICGLLMGILLFLLIQP 492

RESULT 12

US-09-949-016-7542
; Sequence 7542, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7542
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7542

Query Match 17.9%; Score 417.5; DB 2; Length 634;
Best Local Similarity 27.5%; Pred. No. 1.1e-31;

Matches 114; Conservative 90; Mismatches 182; Indels 29; Gaps 9;

QY 2 ARRTPEPDGMRVVVLSAFFQSALVEGVLSFGVFVEFVAAFEQAAARV----SWIAS 57
Db 184 ARGFPPEGGFGWVVFPAATWCNSIFGIHNSVGILYSMLLEEKEKNRQVEFOAAWVGA 243
QY 58 IGIAVQFGSPVGSALSTKFGPRPVMTGILALGMLLASFATSLTHLYLSIGLLSGSG 117
Db 244 LAMGMTPEGSPVIVSIFTDRLGCRITATAGAAVAFIGLHTSFTSSLSLRYFTYGILFGCG 303
QY 118 WALTEAPTLACISCFSSRRSLATGLALTGVGSSFTFAPFQWLLSHYAWRGSLLVSA 177
Db 304 CSFAFQPSLVILGHYFQRRILGLANGVVSAGSSIFSMSFPFLIRMLGDKIKLAQTFQVLST 363
QY 178 LSLHLVACGALLRP--PSLAEDPAVGGPR-----AQLTSLHHPFLRYT-----VAL 223
Db 364 FMFVLMMLSLTYRPLLPSSQDTPSKRGVRTLHQRFQAQLRKYFNNRVERQRTYRIWAFGI 423
QY 224 TLINTGYFTPYLHLVAHLQDLWDPLPAFLLSV-VAISDLVGRVSGWLGDAVPG-PVT 281
Db 424 AAAALGYFVYVHLMKYVEE-EFSEIKETWVLLVCIGATSGGLRVLSGHISDSIPGLKXI 482
QY 282 RLLMLWTTLTGVSALFPVAQAPTALVALVAAYGFTSG---ALAPLAFSVLPBELLGTRR 337
Db 483 YLQVLSFLLGLMSMWIPLCRDFGGLIYVCLFLGLCDGFFITIMAPIAF---ELVGPMD 538
QY 338 IYCGLLQMIESIGLLGPPLSGYLKRDVSGNYTASFVAVAGAFLLSGGILLTLP 392
Db 539 ASQAIGYLLGMALPMIAGPPIAGLLRNCFGDYHVAFYFAGVPPIGAVILFVPV 593

RESULT 13

US-10-074-547-4
; Sequence 4, Application US/10074547
; Patent No. 6989441
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: Millennium Pharmaceuticals Inc.
; FILE REFERENCE: MP12001-019P1RCP1 (M)
; CURRENT APPLICATION NUMBER: US/10/074, 547
; CURRENT FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/269072
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus
US-10-074-547-4

Query Match 17.1%; Score 398.5; DB 2; Length 611;
Best Local Similarity 21.7%; Pred. No. 7.5e-30;
Matches 133; Conservative 81; Mismatches 141; Indels 259; Gaps 14;

QY 20 AFFQSALVEGVLSRFG-VFEVEFVAAFEQAAARVSW--IASIGIAVQFGSPVGSALSTK 76
Db 1 SFLINGFTDGFPKSFGILIFFRELQEEFQASNBETSWDSISSILLAVLFAAGLASILVNR 60
QY 77 FGPRPVMTGILALGMLLASFATSLTHLYLSIGLLSGSGWALTFAPTLACISCFYSRR 136
Db 61 FGCRLVTIAGGLASSGMVLSFATNISELYLTFGVITGLGFAFIYLPAYIYITSYFEKK 120
QY 137 RSLATGLALTGVGSSFTFAPF--FQWLLSHYA--WRGSLLVLS-----ALSLH 181
Db 121 RSLATGIAVAGSGVGFVLAFLPNPDQFLIENYGSKMGALLFFGCMGYIAIWSVAIVLN 180
QY 182 LVACGALLRP-----PSL 194

Db 181 CCIAGALFRPLPSEKVKOTKLAKAEEPKEALKSKENEASESIDSIRSAKAIIVSPETPAL 240
QY 195 AFDPAVCGPRAQITSLHMG-----PFLRY--TVALLTINTGYFI 232
Db 241 SLSPELTPKQDLOKLLKTSRTRSSNGAKLLDFSVLKDARGFLLYASSGSLASTGLFL 300
QY 233 P-----YHLVAH 240
Db 301 PGSIPLVNEAKSLGESLSSVKSKEAFLLSILGDSSDKEGFGIFARPATLLSFLGFVAN 360
QY 241 LQD-----LDWDP LPAF-- 253
Db 361 LKETKSNRPVLIYLLSLCSIVAVVINGILSRLASALAGSRKEKKIKSMIDKIELKSTFWG 420
QY 254 -----L LSVV 258
Db 421 LFLFLSFPFGVGFSGKAVVILALGFLFSILVAIPVVGLOKYSSALGLTETDASTLIEAI 480
QY 259 AISDLVGRVNSGWLGDVAVPGPVTRLMLMTTLTGVSIALFPVQAAPTALVALA----- 311
Db 481 AVLNITGRPLAGLLADKTKNRKLAIVNLSLILCGLFVAFAPLA---TIFGLAFYCVLFG 537
QY 312 ----VAY---GFTSGALAPLAFSVDELIGTRRIYCGILQMIIESIGLLGPPLSGYL 363
Db 538 SIVFLLAYAFKFGCKGSYIALTSVAVDLTGLDKLSNAFGLLLFQGVATLVGPPIAGLL 597
QY 364 RDVSGNYTASFVVA 377
Db 598 KDLTGSYKVSFYFA 611

RESULT 14
US-09-270-767-44472
; Sequence 44472, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44472
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44472

Query Match 16.1%; Score 376; DB 2; Length 480;
Best Local Similarity 37.3%; Pred. No. 8.1e-28;
Matches 69; Conservative 44; Mismatches 72; Indels 0; Gaps 0;
QY 7 PPDGWMGRVVLSAFQSAIVFGVLRSGVFVEFVAAFEEOAARVSWIASIGIAVQFG 66
Db 93 PPDGCGWVICFASFCMNMIVDGIATFGIFLEEFVAYFHEGKTVAWVGSLLSGVYLSA 152
QY 67 SPVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYSIGLLSGSGWALTFAPTL 126
Db 153 GPIVSALANKYGCRAVCIAIGSIACIAFVLSTFSTNVSMLMATYGFMGFGFMILPAV 212
QY 127 ACUSCYFSRRSLATGALTGVGLSSFTFAPFFQWLLSHYAWRGSLLVASLSLHVACG 186
Db 213 VAVGYVFETKRSLATGIAVCGSGFTFAFAPLATYLLSEYGWKNALLIFAGLLNCAIFG 272
QY 187 ALLRP 191
Db 273 AMWRP 277

RESULT 15
US-09-328-352-4392
; Sequence 4392, Application US/09328352

; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER.
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4392
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4392
Query Match 15.0%; Score 350; DB 2; Length 429;
Best Local Similarity 29.0%; Pred. No. 2.3e-25;
Matches 117; Conservative 69; Mismatches 171; Indels 46; Gaps 16;
QY 13 GRVVLSAFQSAIVFGVLRSGVFVEFVAAFEEOAARVSW--IASIGIAVQ--FGS 67
Db 43 GSALIL-----ALSLGVRHGFGLYLVPMSEHF-----GWGQHVFSLAIAQNLWGA 89
QY 68 --PVGSALSTKFGPRPVMTGGILALGMLLASFATSLTHLYSIGLLSGSGWALT-FAP 124
Db 90 IQPFTGAIADKYGSKVVTVGGLLYTLGLLMAFSSSVLILNLSGLIIGLALSATSFTV 149
QY 125 TLACLS-CYFSRRSLATGALTGVGLSFTFAPFFQWLLSHYAWRGSLLVASLSHLV 183
Db 150 LLSAVGRAAPPEKRSMAAGIASAGSFGQFIMLPSTLLLLKTGVW--SSALMVSALLIALI 208
QY 184 ACGA-LLRPPSLAEDPAVGGPRAQITSLH---HGFRLRYTVALTLINTGYFIPLY--H 236
Db 209 IPLAWMLKSPNQTPKAIAQPQLTFKQVLIARKHKEP--WWLALGFLVCGFQVFLGVH 266
QY 237 LVANLQDLWDPLPAFLLSVVAISDLVGRVNSGWLGDVAVPGPVTRLMLMTTLTGVSIA 296
Db 267 LPGYLIDHGFDTTGTVFLALVGLFNIVGTGAGWLGDREKSP--KLMLALYSGRGIAIT 324
QY 297 LFPVAQAPALVALAVAYGFTSGAL---APLAFSVLPDELIGTRRIYCGILQMIIESIG 352
Db 325 AFLLEPLSTYTV--YAFGIMGLMLSTVPLTNGIVANMGVKKYLSMLSGIVFTTHQVG 381
QY 353 GLGPPLSGYLRLDVSGNYTASFVVAAGAFLLSGGILLTPHFF 395
Db 382 SFFGWLGVNHDLTGNYNALWLSIAL-----SLIGVIHFF 419

Search completed: September 7, 2006, 13:08:58
Job time : 52 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2006, 13:02:26 ; Search time 41 Seconds
(without alignments)
1072.465 Million cell updates/sec

Title: US-10-654-428-2
Perfect score: 2331
Sequence: 1 MARTEPPDGMGRVVVLSA.....SLTAPGLLPRLGLHRTVP 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PTR_80:.*
1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	591	25.4	484	2	A55626 monocarboxylate tr
2	564	24.2	542	2	JC5507 monocarboxylate tr
3	555.5	23.8	494	2	A44458 probable transport
4	546	23.4	494	2	JC4399 monocarboxylate tr
5	540	23.2	500	2	A55568 monocarboxylate tr
6	482.5	20.7	598	2	T32430 hypothetical prote
7	463.5	19.9	556	2	T16790 hypothetical prote
8	417.5	17.9	613	2	I39295 X-linked PEST-cont
9	370.5	15.9	397	2	G69295 oxalate/formate an
10	368.5	15.8	402	2	F95874 hypothetical prote
11	368.5	15.8	596	2	T23685 hypothetical prote
12	342	14.7	571	2	T33941 hypothetical prote
13	314.5	13.5	808	2	T16564 hypothetical prote
14	290.5	12.5	399	2	F83484 probable MFS trans
15	288	12.4	410	2	G82445 probable oxalate/f
16	280	12.0	406	2	AG2704 MFS permease (impo
17	280	12.0	414	2	G97486 probable mfs trans
18	250.5	10.7	407	2	AC1346 antiporter protein
19	250.5	10.7	416	2	G69748 conserved hypothet
20	249.5	10.7	405	2	H81995 antiporter protein
21	245.5	10.5	407	2	AF1716 antiporter protein
22	245	10.5	418	2	F83986 transporter BH2694
23	241	10.3	426	2	T50944 hypothetical prote
24	238.5	10.2	402	2	S47768 hypothetical 43K p
25	236	10.1	673	2	S59263 probable membrane
26	233	10.0	400	2	H91182 probable resistanc
27	232.5	10.0	400	2	D86029 probable protein
28	232.5	10.0	405	2	T44249 transport protein
29	227.5	9.8	431	2	AB2949 MFS permease (mono

30	229.5	9.8	433	2	H98333 hypothetical prote
31	228.5	9.8	398	2	AD0828 probable transmemb
32	226.5	9.7	430	2	S74039 hypothetical prote
33	226	9.7	473	2	S38065 hypothetical prote
34	225	9.7	444	2	E83033 probable MFS trans
35	222	9.5	419	2	H83965 hypothetical prote
36	220	9.4	420	2	H69250 oxalate/formate an
37	218	9.4	388	2	AG3307 oxalate/formate an
38	208	8.9	373	2	S00656 hypothetical prote
39	206	8.8	408	2	H95184 oxalate/formate an
40	206	8.8	408	2	H98051 major facilitator,
41	205	8.8	412	2	T37042 probable nitrate/n
42	201	8.6	501	2	S66816 probable membrane
43	199.5	8.6	371	2	B90267 transport protein,
44	199.5	8.6	807	2	T20080 hypothetical prote
45	196.5	8.4	428	2	A75423 multidrug resistanc

ALIGNMENTS

RESULT 1
A55626
monocarboxylate transporter MCT2 - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 18-Sep-1998
C:Accession: A55626
R:Garcia, C.K.; Brown, M.S.; Pathak, R.K.; Goldstein, J.L.
J. Biol. Chem. 270, 1843-1849, 1995
A:Title: cDNA cloning of MCT2, a second monocarboxylate transporter expressed in differ
A:Reference number: A55626; MUID:95130566; PMID:7829520
A:Accession: A55626
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-484 <GAR>
A:Cross-references: UNIPARC:UPI0000178E64; GB:I31957
C:Superfamily: monocarboxylate transporter MCT2
C:Keywords: transmembrane protein

Query Match 25.4%; Score 591; DB 2; Length 484;
Best Local Similarity 33.1%; Pred. No. 9.7e-33;
Matches 135; Conservative 72; Mismatches 175; Indels 26; Gaps 4;

QY	7	PPDGMGRVVVLSA	FFQSA	LVFGVLR	SGVFVEFVA	FEQARVSWIASIGIAVQFG	66
Db	14	PPDGMGWVVGAA	FISIGFSYA	FPKAVTVF	FKDIOQIFQAS	YSETIAWISSIMLAVMYAG	73
QY	67	SPVGSALSTKFGPRPVMTG	LALGMLLAS	FATSLTHLYLSIGLISGSGWALTFAPTL	126		
Db	74	GPISSVLVNNYGSRPV	IIGLLCTG	MTILASFSNSVLELYLTIGFICGLA	FNLQPAL	133	
QY	127	ACLSGYFSRRSLATG	LATGVGLSSFTFAPFQWL	SHVWRGSLLVLSALSLHLVACG	186		
Db	134	TIIGKYFYRRRPMANG	LAMAGSPVFLSSLAPFNQYLFNSY	GWKGSFLILGIFLHSCVAG	193		
QY	187	ALLRP----	PSLAEDPAVGPR	AQLT-----	SLHNGPFLRYTVAL	223	
Db	194	CLMRPVQTS	SPRKS	SKSKYGRQDGS	MMKASKSVTAEKINRFLDFSLFKHGRGFLYLSGN	253	
QY	224	TLINTGYFI	PYLHLVAHLODL	WDPLPAFLSVVAISDLVGRVSGWL	GDA--VGPVPT	281	
Db	254	VIMFLGF	FAPIIFLAPYAKDKGV	DEYNALLSVMAFVDMFARPTGL	IANSKLIRPRIQ	313	
QY	282	RLMLMTLLTG	VSLALFPVAQAPTAL	VALAVAGFTSGALAPLAFSVLP	PELIGTRRIYCG	341	
Db	314	YFFSPAIVGT	GICHLCLPLADTYP	ALVVYSIFFGYGFSVSVLFETLMDLVGP	PARFSSA	373	
QY	342	IGLLQMI	ESIGLLGPPLSGYL	RDVSGVNTAS	FVAVAGAFLLSGS	GIL	389
Db	374	VGLATIVEC	CPVLLGPPLAGKLV	DKTKDYKMYIASGT	IVVI-SGITVL	420	

RESULT 2

RESULT 2

JC5507
monocarboxylate transporter 3 - chicken
C:Species: Gallus gallus (chicken)
C:Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 09-Jul-2004
C:Accession: JC5507
R:Yoon, H.; Fanelli, A.; Grollman, E.F.; Philp, N.J.
Biochem. Biophys. Res. Commun. 234, 90-94, 1997
A:Title: Identification of a unique monocarboxylate transporter (MCT3) in retinal pigment epithelial cells
A:Reference number: JC5507; MUID:97312526; PMID:9168967
A:Accession: JC5507
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-542 <YOO>
A:Cross-references: UNIPROT:Q90632; UNIPARC:UPI0000171286; GB:U15685
A:Experimental source: retinal pigment epithelial cell
C:Comment: This protein regulates lactate levels in the interphotoreceptor space.
C:Superfamily: monocarboxylate transporter MCT2
C:Keywords: phosphoprotein
F:141/Binding site: phosphate (Thr) (covalent) (by casein kinase II) #status predicted
F:210/Binding site: phosphate (Thr) (covalent) (by cAMP- and cGMP-dependent kinases) #status predicted
F:234,467/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted
F:262,526/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 24.2%; Score 564; DB 2; Length 542;
Best local similarity 30.5%; Pred. No. 7.3e-31;
Matches 146; Conservative 82; Mismatches 185; Indels 66; Gaps 9;

QY 6 EPPPYGWRVVVLSAFQOSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGIAVQOF 65
Db 16 KPPDGGWGWIVLFGCFVITGFSYAFPKAVSYVFKELMKDFHVGYSPTAWISSIMLMLYG 75

QY 66 GSPVGSALSTKFGPRPVMTGILALGMLASFATSLTHLYLSIGLSSGWAFTAPT 125
Db 76 TGMCCSIMVNOFCRPVMLIGGLASSGMLASFNTNIELYLTAGVLTGLMALNFQPS 135

QY 126 LACLSQCYFSRRSLATGLALTGVGLSSFTFAFPQWLISHYAWRGSLLVASLSHLVAC 185
Db 136 LIMI GTYFDKRRPLANGIAAAGSPVFLSSPLGQVLEKFGWRGFLIMGILLNCCTC 195

QY 186 GALLRP-----PSLAED-----PAVG-----GPRRA 205
Db 196 GAVMRPLDAGMKRTEKAQDKYEAKEMLPMGKSEEGISTTDGTTKTKAKKKPKGKXL 255

QY 206 QLTSLHGHGFLRYTVALTLINTGYFIPYLHVANLQDLMDPLPAFLLSVVAISDLVG 265
Db 256 LDFSIFSNRGFIITYTISKFILVGLFVPPILVNAYAKDTGVPDTEAFLSIIGFIDIFA 315

QY 266 RVVSG-----WLGDAVPGPVTRLMLMTLTGVSLLAFPVAAQAPFALVALAVAYGFTSG 319
Db 316 RPACGMVAGLKWVRPHV-AYLFSFAMLENGLTDICSAAR--ASNYTGVIIFCVFPGISYG 371

QY 320 ALAPLAFSVPELIGTRRIYCGLLQMTESIGLLGPPLSGYLRDVSNGNYTASFVAGA 379
Db 372 MVGALQFEVLMATVGSQKFSASIGLVLLIEAFVLLIGPPSAGRLVDALKNYEVIFFLAGS 431

QY 380 FLVSGSGILTLPHFPCFSTTSGPQDLVTEALDTKVPPLPKEGLEGLNSTESGPESQS 438
Db 432 EVVL-SALFLAMATYCCCLNRGKKT-----PPEKNPSAGGSDTE--EAES 474

RESULT 3
A44458
probable transporter Mev - Chinese hamster
N:Alternate names: mct1
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
C:Accession: A44458; I48115
R:Kim, C.M.; Goldstein, J.L.; Brown, M.S.
J. Biol. Chem. 267, 2313-2321, 1992
A:Title: cDNA cloning of MEV, a mutant protein that facilitates cellular uptake of mevalonate
A:Reference number: A44458; MUID:93054638; PMID:1429658
A:Accession: A44458
A:Status: preliminary; not confirmed with conceptual translation

A:Molecule type: nucleic acid
A:Residues: 1-494 <KIM>
A:Cross-references: UNIPARC:UPI000017072E; GB:M97382; NID:g191143; PIDN:AAB59630.1; PID:5
A:Experimental source: CHO clone met-18b-2
A:Note: sequence extracted from NCBI backbone (NCBIP:118221)
R:Garcia, C.K.; Goldstein, J.L.; Pathak, R.K.; Anderson, R.G.W.; Brown, M.S.
Cell 76, 865-873, 1994
A:Title: Molecular characterization of a membrane transporter for lactate, pyruvate, and carnitine
A:Reference number: A49951; MUID:94170387; PMID:8124722
A:Accession: I48115
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-359, 'F', 361-494 <RES>
A:Cross-references: UNIPARC:UPI000012F3BF; GB:I25842; NID:g472332; PIDN:AAB59731.1; PID:5
C:Superfamily: monocarboxylate transporter MCT2

Query Match 23.8%; Score 555.5; DB 2; Length 494;
Best local similarity 31.8%; Pred. No. 2.5e-30;
Matches 135; Conservative 64; Mismatches 182; Indels 43; Gaps 5;

QY 7 PPDGWRVVVLSAFQOSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGIAVQOF 66
Db 13 PPDGGMWAVVVGAFISIGFSYAFPKSITVFKEIEGIFNATTSEVSWISSIMLAVMYAG 72

QY 67 SPVGSALSTKFGPRPVMTGILALGMLASFATSLTHLYLSIGLSSGWAFTAPT 126
Db 73 GPISVLVKNKYGRPVMIAGGCLSGCGLIAASFNTVQELYLICIGVIGGLAFNINPAL 132

QY 127 ACLSCYFSRRSLATGLALTGVGLSSFTFAFPQWLISHYAWRGSLLVASLSHLVAC 186
Db 133 TMIGKYFYKKRPLANGIAMAGSPVFLSTLAPLNQAFPGIFGWRGSFLILGILLNCVAG 192

QY 187 ALLRP-----PSLA-----DPAVGPRAOQLTSLH----- 212
Db 193 SLMRPIGPKPKIEKLSKESLOEAGKSEANTDLMGSPKGEKRSVLQTIKFLDLSLA 252

QY 213 HGFPLRYTVALTLINTGYFIPYLHVANLQDLMDPLPAFLLSVVAISDLVGRVVGWL 272
Db 253 HRGFLLYLSGNVMPFGLFTPLVFLSNYGSQHSSEKSAFLLSILAFVDMVAPRSMGLA 312

QY 273 GDA--VPGPVTRLMLMTLTGVSLLAFPVAAQAPFALVALAVAYGFTSGALAPLAFVLP 330
Db 313 ANTKWIRPRIQYFFPAASVANGVCHLLAPLSTYIGFCIYAGVFGFACGWLSSVLFETLM 372

QY 331 ELIGTRRIYCGLLQMTESIGLLGPPLSGYLRDVSNGNYTASF-----VVAQAFLLS 383
Db 373 DIVGPQRFSSAVGLVTIVECCPVLGLPPLGRLNDMWGDKYTYWACGVILIIAGIYLF 432

QY 384 GSGI 387
Db 433 GMGI 436

RESULT 4
JC4399
monocarboxylate transporter - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: JC4399
R:Takanaga, H.; Tamai, I.; Inaba, S.; Sai, Y.; Higashida, H.; Yamamoto, H.; Tsuji, A.
Biochem. Biophys. Res. Commun. 217, 370-377, 1995
A:Title: cDNA cloning and functional characterization of rat intestinal monocarboxylate transporter
A:Reference number: JC4399; MUID:96095680; PMID:8526936
A:Accession: JC4399
A:Molecule type: mRNA
A:Residues: 1-494 <TAK>
A:Cross-references: UNIPROT:P53987; UNIPARC:UPI000012F3C2; GB:D63834; NID:g1199781; PIDN:
A:Experimental source: Small intestinal mucosal cells
C:Genetics:
A:Gene: mct1
C:Superfamily: monocarboxylate transporter MCT2
C:Keywords: intestine; transmembrane protein
F:16-36/Domain: transmembrane #status predicted <TML>

Db 17 PPDGKGVVVLGSFFVHVADGFVYSPFGLVETLMEEFHASNTMAALLISLLTGLTGS 76
QY 67 SPVTSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSGWAALTFAPTL 126
Db 77 GPLASAVCNKYGCRITTTITGAGVAFVGCALSYFATAMWHIVITGVIMGTGFLMYCPAI 136
QY 127 ACLSCYFSRRRLATGLALTGVGLSSFTFAFFQWLLSHYAWRGSLLVLSALSLHLVACG 186
Db 137 TTTMYFENKRSLATGIAVAGAGVTVLFAPINAFINNNGWRSVFLAFLFILVLCALCG 196
QY 187 A-----LRRPPL----- 194
Db 197 ATFAPLEFALVPDEEEVNDVLDVDKKTAAANEHGEKATLLSPPSLERSMSQSSGVQ 256
QY 195 -----AEDP--- 198
Db 257 KAGEKIRPASSMGADVESQPRSRRTGVERDSGYLNRKDVFTGSI SNVAEFREDPDKY 316
QY 199 -----AVGPPAQLT----- 208
Db 317 RSTGSLHGRHTTVGSI PAHSTGRLDVREGSEEDSKAMDIS EKTNGTNDTNDVEGKMEK 376
QY 209 -----SLHNGPFLRYVALTLINTGYFIPYLHLVAHLQDLMDPLPAAFLLSVAI 260
Db 377 TISNMLSLLELLEPTFLFALSNLLTSVGNSPLVFLPLHATKIGLEPIQGSKVLSAFGV 436
QY 261 SDLVGRVNSGWLGD---AVGPPV---TRLLMLW-----TTLTGVSIALFPVAQAPTALVA 309
Db 437 SNTVGRILFGVVADHKLPLPGGIGNDTARNRLWMYNISLITCGLLTVFCYQFNGFIPLAT 496
QY 310 LAVAVGFTSGALAPLAFSVLPBELIGTRRIYCGLGLLQMTESIGLLGPPLSGYLRDYSGN 369
Db 497 YSALFGFSIASYICLTSVILVDLGLDKLTNAFGLLLWQGVGTVEGPPVSGYLADMTGN 556
QY 370 YTASFYVAGA-----FLLSGGILLTLPHFPCF 397
Db 557 YTLSPVTCGVNLLTYFLQFHLSSGFLVHFVHYLGF 592

RESULT 7
T16790
hypothetical protein T02G5.12 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T16790
R;Pauley, A.
submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid T02G5.
A/Reference number: 218577
A/Accession: T16790
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-556 <PAU>
A/Cross-references: UNIPROT:Q02109; UNIPARC:UPI000007FB4C; EMBL:U41105; NID:g1086772; P1
A/Gene: CESP:T02G5.12
A;introns: 50/3; 100/2; 142/2; 184/3; 350/1; 435/3; 528/3

Query Match 19.9%; Score 463.5; DB 2; Length 556;
Best Local Similarity 25.4%; Pred. No. 4.7e-24;
Matches 134; Conservative 76; Mismatches 178; Indels 139; Gaps 9;
QY 7 PPDGKGVVVLGSFFVHVADGFVYSPFGLVETLMEEFHASNTMAALLISLLTGLTGS 66
Db 14 PPDGKGVVVLGSFFVHVADGFVYSPFGLVETLMEEFHASNTMAALLISLLTGLTGS 73
QY 67 SPVTSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSGWAALTFAPTL 126
Db 74 GPLASAVCNKYGCRITTTITGAGVAFVGCALSYFATAMWHIVITGVIMGTGFLMYCPAI 133
QY 127 ACLSCYFSRRRLATGLALTGVGLSSFTFAFFQWLLSHYAWRGSLL----- 174
Db 137 TTTMYFENKRSLATGIAVAGAGVTVLFAPINAFINNNGWRSVFLAFLFILVLCALCG 196

QY 175 -----VSALSLHLV-----ACGA- 187
Db 194 DGEYNDVFEDEVKKIDYNKRFEESKLSLSPQVLEKSMQCCDIQKNQHARPSSSVGA 253
QY 188 -----LRRP-----PSLAE---DF---AVGPPAQLTSLH 212
Db 254 DLEPLRPRCGTIGERNSGYLNRKDVFTGSI SNVAEFPRANPDKYRSTGSLHTALNIHTH 313
QY 213 H-----GPPLRYTVALTLINTGYFIPYLHLV 238
Db 314 HTGTVGDDRESGHGCDQSLDGMNEKETDFMDIEGTFLFAISNLLTSVGNSPLYFLP 373
QY 239 AHLQDLMDPLPAFLLSVVAISDLVGRVNSGWLGD-----AVGPPVTRLLMLW----- 287
Db 374 LHTNITLSDNNVSGNLSVYGTANTIGRLIFGVVADQKLPLPAGLGDDTARNRLWMYNIS 433
QY 288 TTLTGVSIALFPVAQAPTALVALAVAGFTSGALAPLAFSVLPBELIGTRRIYCGLGLLQ 347
Db 434 LTICGIYTAFCYKFDGYILLDIYAVLFGFTIASVCLTSVILVDLGLDKLTNAFGLLL 493
QY 348 IESIGLGPPLSGYLRDVS GNYTASFVAGAFLLSGGILLTLPHF 394
Db 494 WQGVGTVIGPLVSGYLADLTGNVTLSEVFCGINLMVSGMLFVLPYF 540

RESULT 8
I39295
X-linked PEST-containing transporter - human
C/Species: Homo sapiens (man)
C/Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C/Accession: I39295; I38495
R;Lafrèniere, R.G.; Carrel, L.; Willard, H.F.
Hum. Mol. Genet. 3, 1133-1139, 1994
A/Title: A novel transmembrane transporter encoded by the XPCr gene in Xq13.2.
A/Reference number: I38495; MUID:95072579; PMID:7981683
A/Accession: I39295
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-613 <RES>
A/Cross-references: UNIPROT:P36021; UNIPARC:UPI000012F3CF; EMBL:U05321; NID:g458253; PID
A/Accession: I38495
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 16-613 <RE2>
A/Cross-references: UNIPARC:UPI000016A07E; EMBL:U05315; NID:g458246; PIDN:AAB60374.1; P11
A/Gene: XPCr
A;introns: 218/1; 266/2; 416/3; 464/3; 541/1

Query Match 17.9%; Score 417.5; DB 2; Length 613;
Best Local Similarity 27.5%; Pred. No. 6.7e-21;
Matches 114; Conservative 90; Mismatches 182; Indels 29; Gaps 9;
QY 2 ARTEPPDGWGRVVLSAFQOSALVFGVLSFGVFFVEFVAAFEQAAARY---SWIAS 57
Db 163 ARGQOPREGFGWVVFATWCNGSIFGIHNSVGLYSMLLEEKEKNQVEFOAAWGA 222
QY 58 IGIAVQFGSPVGSALSTKFGPRPVMTGGILALGMLASFATSLTHLYLSIGLSSG 117
Db 223 LAMGMIFFCSPIVSIFTDRLGCRITATAGAAVAFIGLHTSSFTSSLSLRYFTYGI LFGCG 282
QY 118 WALTAPTLACSCYFSRRSLATGLALTGVGLSSFTFAFFQWLLSHYAWRGSLLVSA 177
Db 283 CSFAFQPSLVILGHYFORRLGLANGVVSAGSSIFSMSFPFLIRMLGDKIKLAQTFQVLST 342
QY 178 LSLHLVACGALRRP---PSLAEDPAVGPR-----AQLTSLHNGPFLRYT-----VAL 223
Db 343 FMFVLMLLSLTYRPLLPSSQDTPSKRGVRTLHQRF LAQLRKYFNMRVFRQRTYRIWAFGI 402
QY 224 TLINTGYFIPYLHLVAHLQDLMDPLPAAFLLSV-VAISDLVGRVNSGWLGDVPG-PVT 281
Db 403 AAALGYFVPYVHLMKYVEE-EFSEIKETWLLVCIGATSGLGLVSGHISDSIPGLKKI 461

QY 282 RLMLMTTLTGVSLLAFPPAQAPTALVALAVAGFTSG----ALAPLAFSVLPGLIGTRR 337
Db 462 YLQVLSFLLGLMSMMPLCRDFGGLIVCLFLGLCDGFFITIMAPIAF----ELVGPMQ 517
QY 338 IYCGLGLLQWIESIGLLGPPLSGVLRDVSGNNTASFFVAGAFLLSGSGLLTLP 392
Db 518 ASQAIGYLLGMMLPMLAGPPIAGLLRNCFGDYHVAFFYAGVPPIIGAVILFFVP 572

RESULT 9
G69295

oxalate/formate antiporter (oxiT-2) homolog - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Nov-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: G69295
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Nature 390, 354-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Hoase, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: G69295
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-397 <KLE>
A;Cross-references: UNIPROT:O29880; UNIPARC:UPI0000057135; GB:AE001079; GB:AE000782; NID
C;Superfamily: hypothetical protein c0103

Query Match 15.9%; Score 370.5; DB 2; Length 397;
Best Local Similarity 28.9%; Pred. No. 6.5e-18;
Matches 114; Conservative 68; Mismatches 171; Indels 41; Gaps 10;
QY 16 VVLSAFFQSALVFGLRSFGVFEFVAAFEQQAARVSWIASIGIAVQQRGSPVGSALST 75
Db 6 LVTFAGFLMLMLGIVYAWSVFVNPLMTFGMSKTIASLPFSIFLLIPALMMVPAGRTQD 65
QY 76 KFGPRPVNMTGILALGMLLASFATSLTHLY--LSIGLLSGSGWALTFAPTLACLSCY 132
Db 66 RIGPRKXAMLGVLVGLVGLLSGLIESIQSPYWLIFS YGVIAAGCGGLGYACPVPARKW 125
QY 133 FSPRPSLATGLALTGVGLSSFTFAPEFQWLLSHYAWR-----GSLLVSALSLLHVACG 186
Db 126 FPERVGLATGLVVMGFGMSALIPALERILIDTYGISTTFYILGVILLIVA----VFPAAS 181
QY 187 ALLRPPSL----AEDPA-----VGGPRAQLTSLLHGPFLRYTVALT-----LINTG 229
Db 182 LLSNPPELPAVQASNPAAKEVVTGKPELGPSEMLRDYRF--YVLWLSFFFMALAGLMVIG 239
QY 230 YFIPYLHLVAHLQDLWDPLPAFAFLLSVAISDLVGRVSGWLGDAVPGPVTRLMLMTT 289
Db 240 HIAFPY-----AQERGLEPDLAAFAVSISSVANAVGRPGAGALSDKIGRAMTFVLF--L 291
QY 290 LTVSVLALFP-VAQAPTALVALAVAGFTSGALAPLAFSVLPGLIGTRRIYCGELLOMI 348
Db 292 IQGITLIAFPYHALTLITITIVICAAIIGFNYGANFSLFPSATGDFFGTKMLGVNYGLVFTS 351
QY 349 ESIGGLGPPLSGVLRDVSGNNTASFFVAGAFLL 382
Db 352 YGVGGLVGPIMAGYVFDVTGSYEIAFLVAGVAL 385

RESULT 10
F95874

probable transporter protein SMB20272 [imported] - Sinorhizobium meliloti (strain 1021)
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: F95874
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernat
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: F95874
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-402 <KUR>
A;Cross-references: UNIPROT:Q92WR7; UNIPARC:UPI00000CB481; GB:AL591985; PIDN:CAC48662.1;
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abolaj, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB20272
A;Genome: plasmid
C;Superfamily: hypothetical protein c0103

Query Match 15.8%; Score 368.5; DB 2; Length 402;
Best Local Similarity 28.4%; Pred. No. 9e-18;
Matches 110; Conservative 61; Mismatches 200; Indels 17; Gaps 5;
QY 9 DCGMGRVVVLSAFFQSALVFGLRSFGVFEFVAAFEQQAARVSWIASIGIAVQQRGSP 68
Db 3 DAHYRWVIVAAGLLIGCIAIGAMFSLPFLVPISRDTGMSVTVGSSAMTVGFIMALASM 62
QY 69 VGSALSTKFGPRPVMTGILALGMLLASFATSLTHLYLSIGLLSGSGWALTFAPTLAC 128
Db 63 VMGSASDRKGRPSVVLIGSALLASLALSSLVTSLIAFQLVFGVVVGACAAIFAPMMAC 122
QY 129 LSCYFSRRSLATGLALTGVGLSSFTFAPEFQWLLSHYAWRGSLLLVSALSLLHVACAL 188
Db 123 VTGWEDTHRS LAVSLVSAGMGMAPMTMSPLAGWLVTITVDMRTSLQIITAAAVTMIPAL 182
QY 189 L--RPPSLAEDPAVGPRAQLTSLLHGPFLRYTVALTLINTGYFI-----PYLHLVAH 240
Db 183 LLRRPVLLEDTNVAVSASEGPDMSL--GQALRSPQFVYILLTNFFCCATHSGPIFHTVSY 240
QY 241 LQDLWDPLPAFAFLSVVAISDLVGRVSGWLGDAVPGPVTRLMLMTTLTGVSLLAFPPV 300
Db 241 AVSCGIPMAAAVSIYSLGLAGMGGRVAFVGLGDRY--GAKRILVSGLLLQAFGALAYFF 298
QY 301 AQAPTALVALAVAGFTSGALAPLAFSVLPGLIGTRRIYCGGLLOWIESIGLLGPPLS 360
Db 299 VRDLGAFYAAVAFGFYIYAGVMPLYAVIARENFPPLRMGMTVIGGTAMAGSLGMAIGPVAG 358
QY 361 GYLRDVSGNNTASFFVA-----GAFLLS 383
Db 359 GLIYDVFASYGWLVIYAGWGLIGIGAFLLA 386

RESULT 11
T23685

hypothetical protein M03B6.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T23685
R;Matthews, L.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19782
A;Accession: T23685
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-596 <WIL>
A;Cross-references: UNIPROT:Q93896; UNIPARC:UPI0000082778; EMBL:Z78545; PIDN:CAB01766.1;
A;Experimental source: clone M03B6
C;Genetics:
A;Gene: CESP:M03B6.2
A;Map position: X
A;Introns: 36/1; 94/1; 273/3; 320/1; 414/3; 533/1

Query Match 15.8%; Score 368.5; DB 2; Length 596;
Best Local Similarity 20.8%; Pred. No. 1.3e-17;
Matches 125; Conservative 88; Mismatches 185; Indels 203; Gaps 14;

QY 7 PFDGGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGIAVQ-- 64
Db 33 PFDGGYMAIVAASFLLNMIVDGVITVG-----KILQPAWVHSFGISEASAA 80
QY 65 -----FGSPVGSALSTKFGPRPVMTGGILAALGMLLASFATSLTHLYLSIGLL 113
Db 81 MTMSILSGCYFFAGPIASVLCNVFGCRQVALAGSLIALGFLLSAPNIYVLYVSFGLI 140
QY 114 SGSGWALTFAPTLACLSCYFSRRSLATGALATGVGLSFTFAPF--FQWLLSHYAMRGS 171
Db 141 AGAGFGMMYLPFAIVIIISOYFAKKRSVATGIAVCGSGIGTTVFALLNDVWDFVGDWKQF 200
QY 172 LLLVSALSLHLVACGALLRP--PS-----LAEDPAVCGPRAQLTS 209
Db 201 LVYTAAVTTISGLACILLRPLKPSKDQIEKVAKVLDEYEEHKEKVPESPLSKHNPFLS 260
QY 210 LL-----HH-----GPFLLR-----YTVALTLI----- 226
Db 261 SLELHTAGKLENGHHTGSIKSIDAVASKEIEDLNKPLRKDIIFYSGSTTNIAARSRTGM 320
QY 227 -----NTG----- 229
Db 321 TREEAHEHQSLRFRGKVEPQOYLSKADILNAGNDVAIANESKIQEIIKQALKQLVDKE 380
QY 230 -----YFIPYLHLVAHLQDLD-WDPLPAAFLLSVVAISDLVGRVV 268
Db 381 LLFSPSFLLLAFSGTLLTCCFYVPFIYLGHNLDKIEGLTTAEKSFVSLGLVNLIVARIG 440
QY 269 SGWLGDAVPGPVTRLMLMTTLTGVSLLALFPVAQAPTLAVALAVAYGFTSGALAPLAFSV 328
Db 441 CGILADRPVVSALAVANNFALIAAGLATMTVPLTYAYWQFLAFCVPFSGVACFAALRSII 500
QY 329 LPELJGTRRIYCGLLQWIESIGGLGPPLSGYLRDVSGNYTASFVAVAGAFLLSGSGL 388
Db 501 VVELIGLEKLSNAFGILLTFMGVGAVVGPSPLAAMKDYTGNFDTSFYVMGA-----L 552
QY 389 LTLPIYFCFSTTSGPDVLVTEALDTKVPPLPKEGLEGLNSTESGPESQSILTAGLLLP 448
Db 553 MTISGAMC-----IPL-----GGLRKRELARSDDSSDAKKAIELOK 587
QY 449 L 449
Db 588 L 588

RESULT 12

T33941
hypothetical protein C01B4.9 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T33941
R/Smith, A.; Wamsley, P.; Fronick, W.
submitted to the EMBL Data Library, February 1999
A/Description: The sequence of C. elegans cosmid C01B4.
A/Accession: T33941
A/Reference number: 721443
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-571 <SMI>
A/Cross-references: UNIPROT:Q9UAT3; UNIPARC:UPI000007D622; EMBL:AF125952; PIDN:AAD14701.
A/Experimental source: strain Bristol N2; clone C01B4
C/Genetics:
A/Map position: 5
A/Introns: 42/1; 99/1; 317/1; 473/1

Query Match 14.7%; Score 342; DB 2; Length 571;
Best Local Similarity 17.7%; Pred. No. 8e-16;

Matches 118; Conservative 76; Mismatches 194; Indels 124; Gaps 11;

QY 4 RTEPPDGGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGIAVQ 63
Db 36 KVTPPDGGYGVVVAASFLLNMVAVDGVITYTCGKILVPIWADQFGSTSVAGAVISILLTCY 95
QY 64 QFGSPVGSALSTKFGPRPVMTGGILAALGMLLASFATSLTHLYLSIGLLSGSGWALTF 123
Db 96 YLAGPLASSFVNVFGRSVAIAGSILATTAPLISRFASTIWELYLLFGVLGGIGFCMYL 155
QY 124 PTLACLSCYFSRRSLATGALATGVGLSFTFA---PFFQWL-----LSHYAWR 169
Db 156 PSIVILSTYFAKKRSVATGIAVCGSGIGTWFSTINGPVFDYFGKDVGSFMYVLAIAIS 215
QY 170 GSL--LLVSALSL--HLVACG-----LRLPSSLAE-- 196
Db 216 GSLFSLFAPLPKATEHQVKVAKMVRNVEGKPEEPTQRLLEDVRNDLEELNRPCHNADTF 275
QY 197 ---DPAVGGPRAQLTSL----- 211
Db 276 YAGNAPVSRSRNTCSFLGIFLKKIQNCLKFRETIKNRKKLDRKAESAEAHVHATEH 335
QY 212 ---HH-----GPFLLRYTVALTLINTGYFIPYLHLVAHLQD 243
Db 336 HTVHVVKKSFTFKKESLCSVLDKDLLFSPSFMTLAVSGTFTVLSFLVPFVYLAAMKQ 395
QY 244 LDWDELPAAFLLSVAIS--DLVGRVSGWLGDAVPGPVTRLMLMTTLTGVSLLALFPVA 301
Db 396 KNPDFTDALSLPVTLLGAFNIMERIGCGMVADHPKMSALQVSNVATIAGTSMLEFVDFC 455
QY 302 QAPTAVALAVAYGFTSGALAPLAFSVLPPELIGTRRIYCGLLQWIESIGGLGPPLSG 361
Db 456 TELMHYVFCIPFSAGVACFALRSVICVELIGVEKLSNAFGILMVFMGIGAVVGPIAA 515
QY 362 YLRDVSGNYTASFVAVAGAFLLSGSGL-LTLP 392
Db 516 QIKDITGNYDISFYVMG-IIFAFSGVMTIRLP 546

RESULT 13

T16564
hypothetical protein K05B2.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T16564
R/Chissoe, S.
submitted to the EMBL Data Library, June 1995
A/Description: The sequence of C. elegans cosmid K05B2.
A/Reference number: 218536
A/Accession: T16564
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-808 <CHI>
A/Cross-references: UNIPARC:UPI000017BA91; EMBL:U29379; NID:g868192; PID:g868197; PIDN:AF
A/Experimental source: strain Bristol N2
C/Genetics:
A/Map position: 5
A/Introns: 28/2; 40/1; 75/1; 122/3; 241/1; 318/1; 415/3; 500/2; 576/1; 642/1; 684/1; 726/1

Query Match 13.5%; Score 314.5; DB 2; Length 808;
Best Local Similarity 17.7%; Pred. No. 8.2e-14;
Matches 122; Conservative 89; Mismatches 219; Indels 261; Gaps 9;

QY 6 EPPDGGRVVVLSAFFQSALVFGVLRSGVFVEFVAAFEEOAARVSWIASIGIAVQF 65
Db 82 QAPDGGYGVVVAFAAFMSNLVVDGISTAFSEFKQSYKERYQESDGLTVFISLITGYLL 141
QY 66 GSPVGSALSTKFGPRPVMTGGILAALGMLLASFATSLTHLYLSIGLLSGSGWALTFAPT 125
Db 142 VGPLVSALCNKYEQRVVMAGGLISGIAFMIAPASPNIVFMLIYGVWGLGFGMITYLP 201
QY 126 LACLSCYFSRRSLATGALATGVGLSFTFAPFQWLLSHYAWRGSLLVLSALSLHLVAC 185

Db 202 I V V G F Y F D S K R A M A T G I S V A G S G V G T F L V P L I C Q V C I A N I G M Q K T I W L A V M I F S C V F Y 261
QY 186 G A L L R P P S L A ----- E D P A V G G P R ----- 204
Db 262 G Y L Y R P L P M V D L N E I R D Q E M E P L R Q A L S K V E D D E D E A V E S P H P N R S T S A S G A I E D E A V S 321
QY 205 ----- A Q L T S L L H H G ----- 214
Db 322 Q N P E V Q R L R S I L T E G D E G S E T P K S D A H K V R S H T A T R S R K H T M T S N G S Q H D L K L S R G N L S 381
QY 215 ----- 214
Db 382 S D N R L S R V S A R D F N Q S L S K S G S G A S N L S I A M S G V D P K E F Q R P M N R Q D I F Y G S I Q N L K 441
QY 215 ----- P F ----- 216
Db 442 E F K E E G T M A G Y R A S T L S I P R S V G Q A V S Q L N L S R T G S R L G G P G I A E D E E M I E P F V D D G C 501
QY 217 ----- L R Y T V A L T L I N T ----- G Y F I P Y L H L V A H L Q D L D W D P L P 250
Db 502 C K V E L P L P M R N A L S E M I D N L L K N K T M I L C I S N L L G M M G F Y I P I M F L K D L S E S M K L D M S L 561
QY 251 A A F L L S V A I S D L V G R V S G W L G D A V P G P V T R L L M T T L T G V S L A L F P V A Q A P T A L V A L 310
Db 562 A M F L V P I F G V E N T I G R V F E G W L T D Q I L S A L T I N N L S L I V S G L L T L A C P L L T S I A G Q Y F Y 621
QY 311 A V A Y G F ----- T S G A L A P L A F S V L P E L I G T R R I Y C G L G L Q M I E S I G G L G P P 358
Db 622 A I V F G F I C K L F Y F Y Q H L E F A Y I C L T S I V L A D M G L E N L T N S F G L L V A R G I A S L V G P P 681
QY 359 L S G Y L R D V S G N V T A S F V V A G A F L L S G S G I L L T L P H F F C F S T ----- T T S G ----- P Q D L V 408
Db 682 F A G I V V D I T G S Y D T A F Y F G G L V L V A G L I S A T I P ----- F V T K K D A D T A D G M P Q L Q D Q D N V 737
QY 409 T E A I D T K V P L P K E G L E G G L N S T E S G P E S Q S L 439
Db 738 S G K L S V L T E R S E A L T D Y Q R T I Q S M K Q Q H Q L 768

RESULT 14

F83484
probable MFS transporter PA1286 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: F83484
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Iarbig, K.; Lim, A.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: F83484
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-399 <STO>
A/Cross-references: UNIPROT:Q9I458; UNIPARC:UPI000000CS2EE; GB:AE004558; GB:AE004091; NID: A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA1286
C/Superfamily: hypothetical protein c0103

Query Match

Best Local Similarity 12.5%; Score 290.5; DB 2; Length 399;
Matches 114; Conservative 62; Mismatches 176; Indels 49; Gaps 17;

QY 16 V L S A F F Q S A L F G V L R S F G V F V E F V A A F E E Q A R V S W --- I A S I G I A V Q O --- F G S P 68
Db 8 I L L S G A L I L A L S L G V R H G F G L F L A P M S A D F ----- G M G R E V F A F A I A L O N L V W G L A Q P 60
QY 69 V G S A L S T K F G P R P V M T G C I L A L G M L A S F A T S L T H L Y L S I G L L S G S G W A L T - F A P T L A 127
Db 61 F T G A L A D R Y G A A P A V L V G G L L Y A L G L V L M G L S Q S A S G L S L S A G L I G L I G L S G T S F S V I L G 120

QY 128 C L S - C Y F S R R R S L A T G L A L T G V G L S S F T F A P F F Q W L L S H Y A M R G S L L V S A L S L H L V A C G 186
Db 121 A V G R A V P A E Q R S M A M G I S S A G S F G Q F A M L P G T L I G L I G M W S S A L L A L G L V A L I V P L A 180
QY 187 A L L - R P - P S L A E D P A V G G P R A Q L T S L L H G P F L R Y T V A L T I N T G Y F I P Y L - H L V A H L 241
Db 181 G L M K D R F L P P Q G H E Q S L G - E A L R E A C A H S G F W L - - L A L G F F V C G F Q V F I G V H L P A Y L 235
QY 242 Q D L D W D P L P A - - A F L L S V A I S D L V G R V S G W L G D A V P G P V T R L L M L W T T L T G V S L A L F 298
Db 236 V D - - - Q H L P A Q V G T V I A L V G L F N V F G T Y I A G W L G R W S K P - - R L L T A L Y L V R G V I V L F 290
QY 299 - - - P V A Q A P T A L V A L A V A Y G F T S G A L - - - A P L A F S V L P E L I G T R R I Y C G L G L Q M I E S I 351
Db 291 L W L P L S - - - - - V Y S A Y A F G V A M G L L W L S T V P L T N G T V A T L F G V R N L S M L G G I V F L F H Q L 344
QY 352 G L L G P P L S G Y L R D V S G N Y T A S F V V A G A F L L S G S G I L L T L P 392
Db 345 G A F L G W L G G V V Y D R T G S Y - - D L W Q L S T I L L S L A L L N W P 383

RESULT 15

G82445
probable oxalate/formate antiporter VCA0554 [imported] - Vibrio cholerae (strain N16961)
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: G82445
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Winn, M.L.; Dodson, R.J.; Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; R. R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: G82445
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-410 <HEI>
A/Cross-references: UNIPROT:Q9KM35; UNIPARC:UPI000000C3579; GB:AE004386; GB:AE003853; NID: A/Experimental source: serogroup O1, strain N16961; biotype El Tor
C/Genetics:
A/Gene: VCA0554
A/Map position: 2
C/Superfamily: hypothetical protein c0103

Query Match

Best Local Similarity 12.4%; Score 288; DB 2; Length 410;
Matches 107; Conservative 70; Mismatches 167; Indels 54; Gaps 14;

QY 17 V L S A F F Q S A L F G V L R S F G V F V E F V A A F E E Q A R V S --- W I A S I G I A V Q Q F G S P V G S A L 73
Db 10 I L L A G F C I N L C L G I L Y A M S V E N K A L V T D F G W S A D A S S P Y A I A T I A F S V C L L - - V A G I L 66
QY 74 S T K F G P R P V M T G C I L A L G M L A S F A T S L T H L Y L S I G L L S G S G W A L T F A P T A C L S --- 130
Db 67 Q D R M G P R N I L I G T L T L G L M I A S G F A S S V L M N L T F G V M T G A G I G F G Y --- A C L S P S A 122
QY 131 - - - C Y F S R R R S L A T G L A L T G V G L S S F T F A P F F Q W L L S H Y A M R G S L L V S A L S L H L - - - V A C 185
Db 123 M K W F H P S K K G W N G L I A A G F G L A A I Y L A P L T S A L I T H L G I Q T S F M I L G A G V L A I A V P L A C 182
QY 186 G A L L R P P S - - - L A E D P A V G G P R A Q L T S L L - - - - - H H G F L R Y T V A L T - L I N 227
Db 183 - T I N N P A G Y V P A E P K I K A G Q E A K V A K V A N L S W K A M L K T P Q F Y A L M L M Y A L A A S V G L M I 241
QY 228 T G Y F I P Y L H L V A H L Q D L D W D P L P A F L L S V A I S D L V G R V S G W L G D A V P G P V T R L L M L W 287
Db 242 I G N I T N I A S V Q A N L P N - - - - - A V Y L A S I L A I F N S G R V A A G I L S D K I G G - - V R T L L A 292
QY 288 T T I T G V S L A L F P V A Q A P T A L V - - - - - A L A V A Y G F T S G A L A P L A F S V L P E L I G T R R I Y C G L G 343
Db 293 F L I O G G N M V L F A T P D S E P T L I I G T A I A A V G Y G - T L L A V F P - - - S I T A E F Y G L K N Y G T N Y G 348
QY 344 L L Q M I E S I G L L G P P L S G Y L R D V S G N Y T A S F V V A G A F L 381

Thu Sep 7 14:12:16 2006

us-10-654-428-2.rpr

Page 8

```

      :| ||| :| :|
Db 349 VLYTSGICGAICAAVVGYSMTHGGYNLA YTISAAM 386

```

Search completed: September 7, 2006, 13:08:03
Job time : 44 secs